Date: December 13, 2001

From: Ellis F. Unger, M.D., DCTDA, OTRR

Subject: Clinical Review of the June 28, 2000 Amendment to Supplemental BLA 98-0737 - - STN103471

Chiron Corp.; Interferon β-1b (Betaseron®)

Through: Marc Walton, M.D., Ph.D., Chief, Medicine Branch, DCTDA
Karen Weiss, M.D., Director, DCTDA

To: STN103471 File

This document is the Medical Officer Clinical Review for the June 28, 2000 Amendment to Supplemental BLA 98-0737

Sponsor: Chiron Corp.
Product: Interferon β-1b (Betaseron®)
Proposed Indication: treatment of secondary progressive forms of Multiple Sclerosis
Proposed Regimen: 0.25 mg SC QOD, indefinitely
Date of Submission:
Original: June 29, 1998

Information covered in this review:

June 28, 2000	–	major amendment to supplement
June 14, 2001	–	Responses to FDA Complete Review Letter; Combined analyses of North American and European SPMS Studies

Table of Contents

Abbreviations used in this review

BSA	body surface area
CMH	Cochran-Mantel-Haenszel
CNS	central nervous system
ECG	electrocardiogram
EDSS	Expanded Disability Status Scale
EU	European Union
FS	Functional Systems
GEE	General Estimating Equations
GEMS	Global Evaluation of Multiple Sclerosis
IMB	Independent Monitoring Board
ITT	intent-to-treat
IFN	interferon
IFNs	interferons
MADRS	Montgomery and Asberg Depression Rating Scale
MRI	magnetic resonance imaging
mIU	million International Units
MS	Multiple sclerosis
NA	North American
NAB	neutralizing antibodies
NSAID	non-steroidal anti-inflammatory drugs
QID	four times daily
QOL	quality of life
RR	relative risk
RRMS	Relapsing-Remitting Multiple sclerosis
SIP	Sickness Impact Profile
SGOT	Serum glutamic oxaloacetic transaminase
SGPT	Serum glutamic pyruvic transaminase
SPMS	Secondary Progressive Multiple sclerosis
SSC	Study Steering Committee
TID	three times daily
TIW	three times each week

Introduction

Interferon β-1b (Betaseron®) was approved by the FDA for treatment of patients with the relapsing-remitting form of multiple sclerosis (RRMS) in 1993. Phase 4 commitments made at the time of Betaseron's licensure included a commitment to validate magnetic resonance imaging (MRI) endpoints with respect to clinical benefit, and a commitment to study the safety and efficacy of the product in subjects with secondary progressive multiple sclerosis (SPMS).

Chiron, Inc. submitted a supplemental BLA on June 29, 1998, seeking approval to expand the indication section to include the treatment of SPMS. The sponsor was granted Fast Track status on May 1, 1998.

-

The latest proposed labeling claims are:

The focus of the original supplemental BLA documents filed on June 29, 1998 was Study ME 93079, a double-blind placebo-controlled multicenter study to evaluate the safety and efficacy of interferon β-1b, administered subcutaneously to outpatients with SPMS. The study was conducted in the European Union (EU) between September 1994 and March 1998 by ----------with data analysis and interpretation by ---------- The EU study was previously reviewed by CBER, and a complete review letter was issued, dated December 28, 1998.

At the time of supplemental BLA filing, a very similar investigation, Study BL01-3112/3118, was being conducted by ----------n North America (NA). The NA study had been initiated in August 1995, approximately one year after the initiation of Study ME 93079 in Europe. In 1998, with full knowledge of the evidence of efficacy from the EU study, and with knowledge of interim analyses of the NA study, the Independent Monitoring Board (IMB) of the NA study deemed continuation of the study to be "…scientifically sound and ethically appropriate," despite its placebo-controlled design. The willingness of the IMB to continue the placebo-controlled NA study in the face of positive results from the EU study raised the possibility that data from the NA study were inconsistent with the results of the EU study, casting doubt on the estimate of the efficacy of Betaseron in SPMS. Therefore, in the course of reviewing the original supplement, CBER requested and was granted permission to have contact with the IMB of the ongoing NA study, and to review interim data as presented to the IMB.

Based, in part, on interim analyses of the NA study, CBER issued a Complete Review letter to the sponsor, dated June 4, 1999, stating that evidence of the product's efficacy in chronic progressive MS had not been adequately substantiated. The letter requested submission of the results of the NA study upon its completion.

These amendments to the supplemental BLA include the phase 3 data, analyses, and study reports from the NA investigation, Study BL01-3112/3118, with integrated analyses from both the NA and EU studies. Numerous exploratory analyses were conducted of the combined EU and NA datasets, both by the sponsor and by CBER, in order to reach an overall interpretation of the data from the 2 studies.

Notes:

1. For overviews of multiple sclerosis (MS), MS treatment modalities, interferons, interferon β-1b (Betaseron®, Chiron, Inc.), and prior Betaseron clinical studies, see the original sBLA 98-0737 review document, dated May 7, 1999.

2. Summaries and interpretations of the sponsor's submitted reports are shown in Arial font; CBER's comments and exploratory analyses are denoted by Times font.

Protocol BL01-3112/3118

Title: Phase III, double-masked, placebo-controlled study to evaluate the safety and efficacy of two doses of Betaseron in patients with secondary progressive multiple sclerosis

Study Period: August 2, 1995 to November 22, 1999 (early termination based on IMB recommendation)

Funding:----------

Centers: Thirty-five sites within the U.S. (31) and Canada (4)

Study Background

---------submitted two protocols for the evaluation of Betaseron in subjects with secondary progressive MS in mid-1995: Protocol BL01-3112/3118 as well as Protocol 93079. The primary endpoint for both studies was time to progression of disability. (The prior licensure of Betaseron for relapsing-remitting MS was based on a reduction in exacerbations.)

Protocol ME 93079 was intended to support registration of Interferon β-1b in the EU, and was not conducted under formal IND regulations. The 718-patient placebo-controlled study, conducted by----------in the EU, compar ed a single active treatment group (Betaseron 0.25 mg SC QOD) to a placebo group. The results are published in the open literature, and were reviewed separately in the clinical review document for the original supplemental BLA. Protocol BL01-3112 was conducted in NA under IND -------- a ~900-subject multicenter study designed to assess the safety and efficacy of two doses of Betaseron in SPMS.

Objectives

The stated primary study objective was to determine if the administration of Betaseron affects disease progression in SPMS.

Design

This was a double-blind, placebo-controlled study of two dosing regimens of Betaseron, administered to U.S. and Canadian outpatients with secondary progressive MS. There were 2 active treatment arms and 1 placebo arm. Subjects were randomly allocated to every other day (QOD) treatment with placebo, Betaseron 0.25 mg (8 MIU), or Betaseron 0.16 mg/ m^2 (5 MIU/m^2) of body surface area (BSA). The placebo group was further randomized so that half received a fixed 1.0 mL of placebo solution and half received a fraction of 1.0 mL based on BSA. The fixed 8 MIU dose was selected because it is the licensed dose for the treatment of RRMS. The BSA-adjusted dose was selected to investigate the potential relation between body size and dose.

To reduce the frequency and severity of flu-like symptoms and thus minimize the potential for patient unblinding, subjects escalated their doses in weekly increments during the first 3 weeks of treatment and received ibuprofen 1200 mg/day during the first 7 weeks of treatment. The planned treatment duration was 156 weeks. The primary outcome measure and several secondary and tertiary outcome measures were based on the Expanded Disability Status Scale (EDSS) score, a standard scale for measuring neurological impairment in MS patients (see below).

A frequent MRI subgroup was examined to assess the effect of Betaseron on newly active lesions. MRI exams, with and without gadolinium, were to be performed every 4 weeks through

Week 156. Scans were to be centrally interpreted by the----------MS/MRI Group.

Randomization and Blinding

The randomization scheme was generated by -------- ---------- Subjects were randomly assigned to receive 8 MIU Betaseron, 5 MIU/m^2 Betaseron, or placebo in a 1:1:1 ratio using a block-size of six. Each block allocated 2 subjects to fixed-dose Betaseron, 2 to BSA-adjusted Betaseron, 1 to fixed placebo, and 1 to BSA-adjusted placebo. Each site received an adequate number of blocks, based on projected subject recruitment, to ensure sequential subject numbering within each site.

The biostatistician and supporting programmers were the only individuals with access to the randomization codes, with the exception of the drug packaging and labeling department at --------- which also had one copy of the code to use in the labeling and packaging of drug. All sponsor personnel directly involved in the conduct of the study, investigators, and subjects remained blinded to subject treatment assignment throughout the study. The active study drug was indistinguishable from placebo in appearance, smell and color, and labels and packages for active study agent and placebo were indistinguishable. Sealed emergency envelopes, containing the identity of the study agent, were produced by the study biostatistician at ---------- ---------- and stored in a secure area at each site. In an emergency, the treating physician was permitted to unblind the code for a particular subject. The person breaking the code was to document the date, time, and reason for unblinding, and was not to reveal the information to the evaluating physician (see below).

Because side effects associated with Betaseron treatment are distinct and well known, there is the potential for physicians and subjects to surmise treatment assignment based on the presence or absence of side effects and laboratory abnormalities. A number of measures were instituted to protect the integrity of the blind. The primary efficacy endpoint was based upon changes in serial assessment of EDSS score; therefore, all routine EDSS evaluations were performed by designated evaluating physicians, separate from the designated treating physicians. Evaluating physicians lacked access to clinical information, other than the current neurologic case report forms, and were prohibited from speaking to subjects except as necessary to perform standardized neurological evaluations. These physicians also lacked knowledge of prior EDSS and Functional System (FS) scores; subjects were to wear clothing over sites of potential injection reactions. New EDSS evaluating physicians were trained by experienced evaluating physicians. Whenever possible, the same evaluating physician performed all routine neurological assessments for a particular subject until the end of the study.

Separate treating physicians were responsible for overall medical care of subjects, as well as the evaluation, management and recording of exacerbations and adverse events. Treating physicians were prohibited from communicating any information to EDSS evaluating physicians that might lead to unblinding.

Adequacy of blinding was evaluated with a blinding questionnaire, which was completed by treating physicians, evaluating physicians and subjects. The questionnaire was administered at Week 12 and at study completion. The three choices provided were: “Betaseron,” “placebo,” and “don’t know.”

Reviewer's Comment:
An answer of "don't know" enables the respondent to avoid making a best guess; more importantly, respondents who believe they know the treatment assignment may be tempted to respond "don't know," because they believe that a correct answer indicates unblinding and undermines the study, whereas a "don't know" response suggests adequate blinding and strengthens the validity of the study. This potential bias in favor of providing a "don't know" response limits the usefulness of the questionnaire.

Patient Population

The intended patient population was subjects with well-documented SPMS with moderate to severe disability, not wheelchair-bound at study entry. Planned study size was 900 subjects.

Inclusion Criteria

- male and female subjects, age 18-65 inclusive
- clinically definite or laboratory-supported definite MS (Poser criteria) with symptoms for ≥ 2 years
- MS in secondary progressive phase (history of relapsing-remitting disease followed by progressive deterioration sustained for ≥ 6 months – progression not to be solely attack-related)
- Kurtzke Expanded Disability Status Scale score (EDSS score) between 3.0 and 6.5, inclusive (3.0 ≡ moderate disability in one functional system, fully ambulatory; 6.5 ≡ constant bilateral assistance [canes, crutches, braces] required to walk ≈20 meters without resting)
- Increase of ≥ 1 EDSS point within previous 24 months, or a 0.5 point increase if baseline EDSS score was 6.5 (estimated or observed)
- At least one relapse and remission in the disease course
- ALT > within 2.5 X upper limit of normal range
- Serum creatinine ≤ 1.5 mg/dL
- Acceptable bone marrow function (hemoglobin ≥ 9 g/dL, WBC ≥ 2.5 and ≤ 15.0 X 10^9/L, absolute neutrophil count ≥ 0.75 X10^9/L, absolute lymphocyte count ≥ 0.75 X10^9/L, platelet count ≥ 100 X 10^9/L)

Exclusion Criteria

- pregnancy
- substance abuse within 60 days of screening
- uncontrolled clinically significant heart disease (angina pectoris, uncompensated congestive heart failure, dysrhythmias)
- relapse within 60 days of screening, or between screening and study Day 1
- inability to undergo MRI
- active peptic ulcer disease or intolerance to non-steroidal anti-inflammatory drugs (NSAIDs)
- any prior use of: total lymphoid irradiation, any interferon, cladribine, monoclonal antibodies
- use (within 6---onths of screening/ after screening) of: cytotoxic or immunosuppressive therapy, glat--amer acetate, other investigational drugs
- use (within 2 months of screening/ after screening) of corticosteroids or ACTH

Written informed consent was to be obtained from all participating subjects.

Reviewer's Comments: Key differences between the enrollment criteria of the NA and EU studies:
- MS duration was required to be 2 years for entry in the NA study, but only 1 year for the EU study.
- Subjects could be as old as 65 years old in the NA study, 55 in the EU study.

- EDSS criteria were less restrictive in the EU study: subjects could be enrolled with 2 clearly-defined relapses in previous 2 years, without documentation of a 1-point increase in EDSS. Also, a 0.5-point increase was acceptable if the baseline score was 6.0 or 6.5. For the NA study, a 1-point increase in EDSS was required (a 0.5-point increase was acceptable if baseline EDSS was 6.5). It is noteworthy, however, that the pre-study increase in EDSS could be observed or *estimated* in the NA study. Thus, in the NA study, there was the potential to enroll more subjects lacking active disease.
- The EU study did not allow inclusion of subjects with relapse or relapse-related neurological deterioration within 30 days of study entry; a relapse within 60 days of screening excluded subjects in the NA study.

*** The overall effect was to enroll, in the NA study, older subjects with longer duration of MS, who had fewer exacerbations and were more established in the secondary progressive phase of their disease.**

Treatment

Material Source

Chiron Corporation supplied Betaseron, placebo and diluent to ----------There were no changes in the appearance of the study medication or packaging during the course of the study. Study agents were to be transported from the study site to patients' homes in insulated containers. Patients were to store study agents at 2 - 8°C (36 - 46°F). Prior to injection, lyophilized powder was to be reconstituted with 1.2 mL of the supplied diluent. The reconstituted product contained no preservatives, and was to be administered within 3 hours.

Dose and Administration

Subjects in the 8 MIU fixed dose group began dosing with 0.25 mL (2 MIU) of study drug on Study Day 1. These subjects increased their study drug dose by 0.25 mL weekly until the dose of 1.0 mL (8 MIU) was achieved (beginning at the 12^{th} dose). Thus, beginning on study Day 23, these patients received 1 mL of Betaseron (8 MIU) SC QOD and continued at this level for the remainder of the study. Subjects in the matching placebo group followed the same dosing paradigm.

Subjects in the 5 MIU/m^2 dose-adjusted group were dosed according to BSA, based on an escalation schedule. These subjects increased their dose weekly until a dose of ~5 MIU/m^2 was achieved (0.5 to 1.0 mL depending on BSA). Patients continued at this level for the remainder of the study. A maximum dose of 12 MIU (1.0 mL) was allowed for patients in the 5 MIU/m^2 group. Dosing of subjects in the matching placebo group followed the same paradigm based on BSA.

Fifty percent (50%) dose reductions were allowed for possible treatment-related AEs or laboratory abnormalities, with resumption of full dose when medically appropriate. Treatment was to be discontinued for grade ≥ 3 toxicity. Exceptions to this were made only if agreed to by the investigator and the trial manager. Exceptions were allowed specifically for lymphopenia. When the toxicity level had decreased to grade 2 or less, therapy was to be reinstituted at 50%, and subsequently escalated to full-dose after consultations between the trial manager and investigator.

Concomitant Medications

To minimize subject unblinding, subjects were to take ibuprofen (400 mg TID) beginning on Day 1 and continuing through Week 7 for prophylaxis against interferon-related side effects. Thereafter, ibuprofen was to be taken as needed. For subjects intolerant to ibuprofen, acetaminophen could be substituted. Antacids and other medications were to be taken as necessary concurrently with ibuprofen for relief of gastrointestinal symptoms.

Systemic Steroids: Systemic glucocorticoids could be administered for the treatment of relapses, but were to be restricted according to a protocol-specified treatment schedule. The only regimen allowed was IV methylprednisolone 1000 mg daily for 5 days. Systemic steroids were to be limited to a maximum of 3 courses within 12 study months, if possible.

Use of cytotoxic, immunosuppressive (other than systemic steroid therapy), immunomodulating, or investigational therapies was not permitted.

Patient Diaries

Patients were supplied with diaries in which they were to record the date and amount of each dose of study agent administered. Missed doses and dose reductions were also to be recorded, along with the reason for the missed or reduced dose. Diaries were to be brought to each scheduled visit and pertinent data were to be transcribed to the CRFs. Diaries were to be maintained as part of patient source records. The diaries were also to be used to collect AEs, intercurrent medical events, and changes in concomitant medications.

Evaluations

Schedule of Evaluations: Non-MRI Substudy Subjects

- Screening/Baseline:
  History, MS history, reproductive/menstrual history*, electrocardiogram, chest x-ray, urine pregnancy test (electrocardiogram and chest x-ray repeated at end-of-study)
- Screening/Baseline, Week 4, Week 12, then every 12 weeks:
  General physical examination, neurological exam (used to produce the FS and EDSS scores), clinical laboratories (hematology, chemistries, urinalysis), neutralizing antibodies, AEs, concomitant therapy, Beck Depression Inventory (the latter also when a subject discontinued study agent)
- Baseline, then every 24 weeks:
  Quality of Life (administered centrally by telephone with staff from University of Toronto) lipid profile, thyroid function studies
- Screening/Baseline, Weeks 48, 96, 156:
  Neuropsychologic assessment
  Environmental Status Scale (ESS)
  Magnetic Resonance Imaging (MRI), both T1 and T2 weighted images
  Serum pregnancy test
- 12 and 156 weeks (or termination):
  Global MS Status Survey
  Blinding questionnaire

*Protocol Amendment III deleted the requirement for collecting data on reproductive and menstrual history.

Schedule of Evaluations: MRI Substudy Subjects

Evaluations were to be performed for the study as a whole; however, subjects in the frequent MRI subgroup were to undergo MRI (with and without gadolinium), neurological examination (used to produce the FS and EDSS scores), and assessment of AEs at baseline, then every 4 weeks.

Definitions and Procedures:

EDSS Scores

The EDSS is a 10 point scale from 0 (no MS symptoms) to 10 (death due to MS). The scale progresses in a single one-point step from 0 to 1, and then from 1 to 10 in 0.5 point increments. EDSS steps through 4.5 indicate no impairment in ambulation; steps 5.0 to 9.5 represent progressive impairment in ambulation and mobility. The EDSS scores were based on standardized neurological evaluations performed by evaluating physicians. Data from the neurological evaluations were used to determine Kurtzke Functional Systems (FS) assessments. The FS scores assess function within individual neurological systems, including visual, pyramidal, cerebellar, brainstem, sensory, bowel and bladder, cerebral and 'other.' The final EDSS scores were based on the combination of FS scores and the subjects' level of mobility. The 'other' portion of the FS assessments was not utilized in the determination of EDSS in this study.

Relapse Assessments

A relapse was defined as the appearance of a new neurologic abnormality or the reappearance of a neurologic abnormality that was present for at least 48 hours and not due solely to fever, known infection, or systemic steroid treatment withdrawal. Subjects were instructed to contact the study center within 72 hours of initially suspecting a relapse. Relapses were considered evaluable only if verified by the treating physician. Such assessments included notation of the functional system(s) involved and a neurological examination comprised of FS, EDSS, and Scripps evaluations. The treating physician noted the start and stop dates of the relapse. Relapses lasting longer than 180 days were assigned a duration of 180 days.

Relapse severity was assessed both objectively and subjectively. For the objective assessment, the Scripps score was used by the sponsor to determine if the attack was mild, moderate, or severe. Mild relapses included those with a decrease of 7 or fewer Scripps points, moderate relapses included those with a decrease of 8 to 14 Scripps points, and severe relapses included those with a decrease of 15 or more Scripps points. For the subjective portion of the evaluation, the treating physician estimated the overall worst severity of the relapse.

Magnetic Resonance Imaging (MRI)

All MRI studies were centrally interpreted at the University of British Columbia, with assessments of lesion activity and quantitative burden of disease. A minimum of 14 days was required between the end of systemic steroid therapy and MRI scanning with gadolinium.

Neuropsychologic Assessment

A psychologist or psychometrician at each site administered the neuropsychologic assessment to all subjects. A normal control group, matched for age, years of education and gender, was included to control for practice-related learning effects encountered in the testing battery. To

ensure uniform application of the test battery, training workshops for the neuropsychologic testing were provided at investigator meetings.

Environmental Status Scale (ESS)

The ESS is a measure of the impact of MS on social roles and employment. The same psychologist or psychometrician who administered the neuropsychologic assessment also administered the ESS. Data are collected, through interview, in 7 areas that include work status, economic status, residential situation, requirements for personal assistance, special transportation needs, requirements for community service and social activities. Each section requires a 0 to 5-point response. Higher point totals indicate progressive dysfunction and impairment.

Removal of Patients from Treatment or Assessment

Treatment could be discontinued in the event of: intolerable AEs, patient decision to discontinue treatment, pregnancy, or use of other investigational cytotoxic, immunosuppressive (other than steroids), or immunomodulating therapies. Subjects who prematurely discontinued study agent were to complete the BDI at their next scheduled visit, and continue routine visits for the duration of the study, unless they were lost to follow-up. Some subjects in the MRI substudy who discontinued study drug early were willing to continue their MRI schedule; others were not, and these patients followed the annual MRI patient visit schedule.

Efficacy Endpoints

Primary Endpoint

The prospectively-defined primary endpoint was time to neurological deterioration. "Time" was defined as the number of days from start of dosing to deterioration; "deterioration" was defined as a one point increase on the EDSS scale from baseline. For subjects with a baseline EDSS of 6 or greater, a 0.5-point increase was considered equivalent to a 1-point increase.

Transient increases in EDSS were not considered as evidence of deterioration. To be considered valid, EDSS increases had to be maintained and confirmed at two subsequent, consecutive 12-week visits. Progression that occurred at the penultimate visit required confirmation by only one subsequent examination, at least 70 days later, but prior to the final study visit. Missing EDSS scores not interpolated, such that EDSS increases, followed by missing scores, required confirmation by two subsequent 12-week scores. Subjects lost to follow-up were censored at the time of loss.

The definition of progression differed somewhat from that of the EU study, and the sponsor provided complementary analyses of progression based on the definition used in the EU study. There were 3 key differences, in that for the EU study: 1) EDSS increases had to be maintained and confirmed at only one subsequent and consecutive scheduled study visit (at least 70 days later), and maintained at any intercurrent visits; 2) EDSS scores obtained during relapses and at unscheduled visits were excluded from calculation of the endpoint. (Relapses for which a date of resolution was not recorded were considered to have a duration of 180 days.); 3) missing EDSS scores were interpolated as the lower of the two bracketing values, assuring that a missing score could not confirm an increase in EDSS.

Secondary Endpoints

There were 5 secondary endpoints:

- Change in EDSS Score – Baseline to Endpoint (baseline ≡ mean of baseline and screening scores; endpoint score ≡ mean of last 2 scores). Changes from 6 to 6.5 and 6.5 to 7 were treated as full 1-point changes.

- Annual Relapse Rate – The number of relapses, confirmed by investigators, divided by time-on-study for each subject.

- Absolute Change in T2 Lesion Volume From Baseline to Endpoint

- Annual Newly Active Lesion Rate

- Change in Composite Neuropsychologic Score From Baseline to Endpoint

Tertiary Endpoints

- EDSS Endpoints
    - proportion of subjects with confirmed disease progression

- Relapse Endpoints – (Only investigator-verified relapses were included in the analyses, i.e., the subject must have been seen while the relapse was ongoing.)

    - annual relapse rate (relapses divided by time on study)
    - time to first relapse
    - number of relapses per subject
    - distribution of relapse severity

        Relapse severity ≡ Δ Scripps score (before – during) relapse. Decreases of 0-7, 8-14 and ≥15 points were considered mild, moderate and severe, respectively. Both Scripps-defined and investigator-defined severities were analyzed, and concordance between the two was assessed.
    - proportion of subjects with moderate or severe relapses
    - annual rates of moderate and severe relapses
    - mean duration of relapses per subject (onset to end; relapses > 180 days were assigned a duration of 180 days)
    - number of days per patient spent in a relapse
    - number of days per patient spent in a moderate or severe relapse
    - number of systemic corticosteroid courses per subject
    - proportion of subjects with systemic corticosteroid use
    - number of MS-related hospitalizations per subject
    - proportion of subjects with time in intensive care unit

- MRI Endpoints
    - Absolute and % change in annual T2 lesion area, baseline to last available on-study scan.
    - Annual rate of newly active lesions (unique lesions that are identified one or more times but which are not in the previous scan)

- Annual rate of new lesions (lesions which appear on the current T2 scan but are not visible on any previous T2 scan)
- Annual rate of newly enlarging lesions (lesions which appear enlarged on the current T2 scan but are stable on the previous T2 scan)
- Annual rate of persistently enlarging lesions (lesions which are categorized as new or enlarging on the previous T2 scan and appear further enlarged on the current T2 scan)
- Annual rate of newly enhancing lesions (lesions which are enhanced on the current T1 scan and are not classified as newly enhancing in a previous T1 scan)
- Annual rate of persistently enhancing lesions (lesions which are enhanced on the current and previous T1 scans)
- Annual rate of persistently active lesions (unique lesions identified one or more times that are also in the previous scan)
- Proportion of newly active scans per subject
- Proportion of subjects with any newly active scans
- Proportion of enhancing scans per subject (defined as any scan with either newly or persistently enhancing activity)
- Proportion of subjects with any enhancing scan

In the lesion area analyses, only subjects with a valid baseline scan and at least one valid on-study scan were included. No correction for missing values was made.

In the frequent MRI scanning cohort, the annualized newly active lesion rate was calculated by computing the cumulative number of newly active lesions and dividing by the number of scans used in the cumulative computation. This resulted in the average number of newly active lesions per scan which was then multiplied by 13 (the number of expected scans per year) to calculate the annual activity rate. Baseline activity was computed as the average of the newly active lesions at screen and Day 1. No correction for missing values was made.

The annual rates of the other types of lesion activity were calculated as described above for the annual rate of newly active lesions and also could only be derived for the frequent MRI scanning cohort. Included as MRI variables in the protocol were recurrent lesion activity and area of non-enhancing lesion volume (derived from the post-gadolinium T1-weighted scan). These data are not available due to technical difficulties encountered by the central MRI analysis center.

- Cognitive function endpoints (change from baseline for the following test variables):
  - selective reminding test
  - 10/36 spatial recall test
  - paced auditory serial addition test (PASAT)
  - symbol digit modalities test
  - word list generation
  - change in composite score, baseline to Months 12, 24, 36 and endpoint
  - number of failed tests ($\leq 5^{th}$ percentile) per subject.
  - Purdue pegboard
  - Trial making tests

  (The neuropsychologic composite index, a 2° endpoint, was computed from these tests)

- Quality of Life: The Multiple Sclerosis Quality of Life Inventory (MSQLI) interviews assessed the following 9 MS-specific domains: bowel and bladder, fatigue, pain-disturbing sensation, physical disability as assessed by the Sickness Impact Profile (SIP) subscales (ambulation,

mobility, body care and movement), visual impairment, sexuality, and perceived cognitive deficits. Global quality of life was assessed by the physical and mental components of the SF-36 survey and a composite index derived from the nine different MS-specific domain scores used in the MSQLI.

- change in score from baseline to each 6-minthly visit for the 9 MS domain scores and the two SF-36 component scores.
- change in score from baseline to each 6-monthly visit for the MSQLI composite score
- environmental status score (ESS) (assessment of social handicap from MS)

- Other Tertiary Endpoints:
  - New use of antidepressants
  - Global impression by subjects and physicians of change in MS

Correlations between MRI parameters and clinical variables

Prompted by communications between the sponsor and CBER in late 1997, these analyses were planned to evaluate the correlation of MRI parameters with clinical outcome as part of the sponsor's phase 4 commitments from the 1993 marketing approval.

Primary correlations:

- *EDSS* – change from baseline EDSS and percent lesion volume (baseline to last scan)
- *Relapse Rate* – annual relapse rate and rate of enhancing lesions activity (newly and persistently enhancing, combined)

Secondary correlations:

- Annual relapse rate and proportion of scans with enhancing activity
- Annual relapse rate and change from baseline to endpoint in lesion area
- Annual relapse rate and annual rate of newly enhancing lesions
- Change in lesion area and lesion activity
- Lesion area and neuropsychologic assessments (baseline, endpoint, and change)

Non-MRI Correlations

- EDSS scores and Scripps scores (baseline, endpoint, and change)
- EDSS scores and quality of life measures (baseline, endpoint, and change)

Safety Endpoints

The safety population was to include all subjects who received at least one administration of the randomized study agent, through 30 days after discontinuation of the study agent. Safety was assessed using the following parameters: Adverse events (AEs), serious adverse events (SAEs), laboratory evaluations, vital signs, EKG, physical exams, Beck Depression Inventory (BDI), and development of neutralizing antibodies to Betaseron.

Adverse Events

Adverse events (AEs) were defined as any illness, sign, symptom, or unfavorable change in clinical status, independent of perceived relatedness to study agent. All symptoms, other than symptoms unchanged from baseline, were to be documented in the Adverse Event CRFs, irrespective of their perceived relation to study drug or coexisting disease. Subject diaries were used to assist in the collection of AEs. For each reported AE, the treating physician recorded the date of onset, date of resolution, severity (mild, moderate, or severe), relationship to study drug (unrelated, unlikely, possible, probable, or definite), action taken with study drug (none, dose reduced, drug interrupted, drug discontinued), AE management (none, medication, other),

whether considered serious, and if serious, whether the subject was hospitalized. The sponsor subsequently coded the AEs using the Hoechst Adverse Reaction Terminology System (HARTS). The severity used was the worst reported severity for each preferred term for each subject.

Selected AEs were analyzed by time to first occurrence using standard life-table methodology. Incidence rates for 3-monthly time intervals were also computed. Time interval incidence was calculated by dividing the number of subjects with the event during the interval by the number of subjects at risk during the interval. A subject experiencing an event that spanned multiple time intervals contributed to the event incidence in each period spanned by the event.

Unscheduled clinic visits were permitted for AEs, disease-related symptoms, or whenever the investigator judged that a visit would be useful to evaluate subject status. If an assessment during an unscheduled visit suggested a relapse or AE, appropriate data were collected via CRF.

Serious Adverse Events

Serious adverse events (SAEs) were defined as events that were fatal or life-threatening, permanently disabling, or required in-subject hospitalization, as well as events that were a congenital anomaly, cancer, or overdose.

Laboratory Evaluations

Laboratory variables were analyzed by comparing changes in values from baseline to various time points. Individual laboratory values were also classified as notable according to parameter values set during the monitoring of the study. The proportion of subjects with toxicity, by worst toxicity per subject, was computed for those parameters assigned laboratory toxicity grades by the protocol.

Vital Signs

Vital signs were assessed as the change from baseline to various time points. Blood pressure was considered notable if the systolic or diastolic pressures were $\geq$ 150 and 100 mmHg, respectively. Temperatures $\geq$ 40°C were considered notable.

Concomitant Medications

Concomitant medications were categorized according to assigned drug class, as well as according to a sponsor-developed dictionary. Medications use was categorized and tallied by treatment group. The number of patients newly treated with antidepressants was also summarized by treatment group. A patient must have used the antidepressant for at least 14 days to be included.

Beck Depression Inventory (BDI)

Depression occurs commonly in subjects with MS, and suicide occurs far more frequently in this group than in the general population. Assessments of depression and mood were performed using the Beck Depression Inventory (BDI), in addition to routine monitoring of depression via AE collection. The BDI is a self-administered, 21-item scale, with each item counting 3 points (the highest possible point total is 63, which represents severe depression). Scoring was as follows: 0-9 $\equiv$ minimal depression; 10-16 $\equiv$ mild depression; 17-29 $\equiv$ moderate depression; and 30-63 $\equiv$ severe depression. A score of 2 or 3 on items 2 or 9 of the BDI is predictive of suicidal

behavior. At the discretion of the treating physician, subjects who reported these scores were referred to the appropriate provider for psychological/psychiatric evaluation and care.

Treatment groups were compared with respect to the following BDI variables: change in BDI score, change in BDI status (minimal, mild, moderate, severe), proportion of subjects classified as moderate or severely depressed on study, proportion of subjects with scores indicating suicidal risk.

Neutralizing Antibodies

Subjects were considered to be NAB positive if they had quantifiable NAB titers (a titer ≥ 20 IU/mL) in serum collected at two consecutive visits. Exploratory analyses were performed using additional titer cutoffs of 100 IU/mL and 400 IU/mL.

Interim Analyses

There were seven interim analyses, approximately 6 months apart, prior to the final analysis. Interim data reports were summarized by treatment group, with groups anonymously labeled. Only the Berlex biostatistician and the programmer producing the tables had access to detailed data.

Reviewer's Comment: The IMB viewed safety data by group. Consideration of data regarding the incidence of injection site reactions and/or leukopenia would have unblinded the IMB to the identity of the placebo group. In light of the study results, however (see Results, below), such unblinding would not be expected to importantly influence the study results.

Statistical Analyses

Primary Endpoint

Time to Confirmed Progression

The Kaplan-Meier product-limit approach was used to describe the distribution of time to progression, and the logrank test was used to test for differences between treatment groups. As supportive statistical tests of the effect of site and baseline disease characteristics on progression, two Cox proportional hazards regression models were used. One was stratified for study site, and the second was adjusted for baseline EDSS and disease duration while controlling for study site.

Secondary Endpoints

Change in Mean EDSS Score (baseline to endpoint)

An ANOVA was used for this analysis; the main effects in the model were treatment, study site, and the treatment-by-study site interaction. A supportive repeated measures analysis of the change in EDSS over time was performed using generalized estimating equations (GEE) methodology. The model included components for treatment and time. The change in EDSS was assumed to be normally distributed and an independent correlation structure was used in the model.

Annual Relapse Rate

A square-root transformation of the annual relapse rates (protocol amendment IV) was to be used to yield normally-distributed data, followed by an analysis of covariance (ANCOVA) to detect treatment group differences.

Absolute Change in Annual T2 Lesion Area (baseline to endpoint)

Treatment groups were compared for absolute (and percent) change in T2 lesion area from baseline to last scan using an ANOVA model based on ranked data and adjusted for the ranks of the baseline lesion burden. The main effects in the model were treatment, study site, and the treatment-by-study site interaction. After examination of the data and performing ANCOVAs based on both actual and ranked data, it was determined that the model based on the ranked data was the most appropriate. A t-test was used to test within-group change from baseline.

Annual newly active lesion rate

To assess the effect of interferon on new lesion formation, a frequent MRI subgroup was to undergo exams before and after administration of a gadolinium contrast agent every 4 weeks throughout the study. Newly active lesions were defined as unique lesions that are identified one or more times but which were not in the previous scan. The annualized newly active lesion rate was calculated by computing the cumulative number of newly active lesions and dividing by the number of scans. The average number of newly active lesions per scan which was multiplied by 13 (the number of expected scans per year) to calculate the annual activity rate. Baseline activity was computed as the average of the newly active lesions at screen and Day 1. No correction was made for missing values.

Treatment group comparisons of the annual rate of newly active lesions were performed using a t-test of the natural log-transformed means. Inspection of the data showed, however, that assumptions necessary for the ANCOVA were not met. Thus, the treatment group comparisons of the proportion of scans per patient that showed new lesion activity and the proportion of scans that showed any enhancement (either new or persistent) were performed using an ANCOVA adjusted for baseline lesion activity. The main effects in the model were treatment, study site, and treatment-by-study site interaction.

Change in composite neuropsychologic score from baseline to endpoint

Treatment group differences were assessed using an ANOVA examining the score for each of the individual neuropsychologic tests and the composite index. The ANOVA used treatment, study site, and treatment-by-site interaction as effects in the model. Treatment group differences in the number of failed tests per patient was assessed using a CMH ANOVA test.

Safety analyses

Treatment group comparisons for the following events were made using a two-sided Fisher's exact test: adverse event incidence, concomitant medication use, grade 3 or greater laboratory toxicity incidence, notable laboratory values, and notable vital signs. Comparisons were pairwise, except in the case of concomitant medication use for which three-group comparisons were performed. For events with very low incidence, no statistical analysis was performed. Total BDI score and changes in laboratory values, vital signs, and BDI score were compared between groups using a Wilcoxon rank sum test. Finally, treatment group differences in the time to first occurrence of specific adverse events were tested using a logrank test.

Evidence of a possible association between neutralizing activity and efficacy was assessed using a Cox regression model to examine the relationship of neutralizing activity to time to confirmed progression, treating neutralizing antibody status as a time-varying covariate. The effect of NAB activity on the secondary efficacy endpoints was explored using general estimating equations methodology.

Treatment differences for change in BDI status were assessed using a CMH ANOVA test to assess the number of categories shifted. Differences in the proportions of subjects moderately or severely depressed or with suicidal risk were assessed using a CMH test for general association. To assess the effect of time on BDI score, a repeated measures ANOVA was used with site, treatment group, site by treatment group, time, and time by treatment group in the model.

The protocol specified the use of an ANOVA model to assess changes in laboratory values, vital signs, BDI score and total BDI score. Due to the early termination of the study, the significance of some of the treatment group differences for most all of the late visits could not be computed. For this reason, the Wilcoxon test was used.

Significance Testing, Allocation of Alpha

All tests were two-sided, at an overall significance level of $\alpha = 0.05$. To adjust for the interim efficacy analyses by the IMB, the bounds proposed by ----------were used. There were seven interim analyses, approximately 6 months apart, prior to the final analysis. A total of 10% of the critical region was allotted for the interim analyses. Therefore, for the primary endpoint analysis, efficacy was tested at $\alpha = 0.0481$.

A two-stage testing procedure was used for the hypothesis testing. In the first stage, the two Betaseron groups were combined for a test against placebo at the required alpha-level for the analysis. If the null hypothesis was rejected at the first stage, the two doses of Betaseron were individually compared to placebo at the same alpha-level. All secondary and tertiary endpoints were considered to be significant at $\alpha = 0.05$. No adjustments were made for testing multiple efficacy variables.

Definitions of Subject Populations

Primary efficacy population. All efficacy analyses were based on an intent-to-treat (ITT) dataset, including all data from all randomized subjects collected through Week 156 or loss to follow-up.
Evaluable efficacy population. A confirmatory analysis was to have been based on an evaluable subset of subjects, with extensive protocol-defined criteria for evaluability. A sponsor committee, blinded to treatment assignment, was to review the protocol-defined criteria for evaluability and to refine these criteria if deemed necessary. According to analyses performed prior to breaking the study blind, this would have led to a substantial (>300) number of subjects with excluded data. Thus, the committee decided that the evaluable efficacy dataset should consist of all data for all randomized subjects up to the quarterly visit after the time the subject completed dosing, and that only the primary endpoint would be reanalyzed using this evaluable subset.
Safety population. Safety analyses include all data collected up to 30 days after study agent discontinuation for subjects who received at least one dose of their randomized treatment.

Study Performance

Study Administration

The study commenced in August 1995. The trial was guided and supervised by a Trial Manager, who was also a member of the Study Steering Committee (SSC). The SSC included a----------biostatistician and selected principal investigators. The SSC met annually and provided expertise and recommendations with regard to study conduct. All members of the SSC remained blinded to treatment assignments throughout the study.

An Independent Monitoring Board (IMB), comprised of experts in neurology or biostatistics who were not directly participating in the conduct of the study, reviewed safety and efficacy data at 6-month intervals. The IMB provided guidance to----------on the basis of their interpretations of these analyses.

At a scheduled IMB meeting in August, 1999, the IMB recommended early study termination after reviewing the available data related to the primary outcome measure, available secondary outcome measures, and safety data. Their rationale was that the primary outcome was unlikely to change meaningfully, on the basis of the limited outstanding data planned to be collected. The recommendation for early study termination was supported by the SSC, and the study was terminated.

Protocol Modifications

The protocol underwent four revisions, summarized below:

Amendment I (June 11, 1995) allowed subjects with an EDSS of 6.5 and subjects up to age 65 (inclusive) to enter the study, altered other inclusion and exclusion criteria, and modified the schedule of activities performed during the study. It also further defined and clarified various tests performed throughout the study and modified the standard steroid regimen used for the treatment of relapses.

Amendment II (October 31, 1995) added the Multiple Sclerosis Quality of Life Inventory (MSQLI) interview to the protocol and eliminated the peripheral category from the clinical toxicity scale. The amendment made minor modifications to inclusion and exclusion criteria and defined and clarified various tests performed throughout the study.

Amendment III (July 10, 1996) incorporated a new statistical section, modified the standard steroid regimen, changed the parameters for grade 2 and 3 granulocyte toxicity, and incorporated the MRI protocol as an appendix. Subjects with a 0.5 EDSS point change from 6.0 to 6.5 in the 2 years prior to screen were permitted into the study per this amendment. The requirement for visits at Day 3 and at Weeks 2 and 3 was removed. The requirement for collecting data on reproductive and menstrual history and menstrual cycle regularity was deleted. This amendment also allowed centers participating in the pharmacokinetic substudy to enroll subjects into the trial who were not participating in the pharmacokinetic substudy.

Amendment IV (August 19, 1997) deleted the requirement for discontinuing study dosing in the event a grade 3 or 4 toxicity recurs during the study. This amendment also made changes to the statistical methodology for analyses of relapse rates. Based on the assumption that relapse rates follow a Poisson distribution, a square-root transformation of relapse rates was added. Amendment IV also eliminated analyses of absolute and percent change in T2 MRI scans in the frequent MRI substudy.

Study Subjects

Patient Enrollment

The first subject was enrolled at the Blake Neurology Professional, LLC, Colorado site on August 29, 1995; enrollment commenced at all sites within an 19-month period, with the last patient enrolled on April 1, 1997. The final patient visit occurred on November 15, 1999. Nine hundred thirty-nine (939) subjects were entered into the study and randomized to treatment. In total, there were 821 subjects enrolled at U.S. sites, and 118 subjects enrolled at Canadian sites.

All 939 randomized subjects received at least one dose of the study agent and comprise the ITT population of the study. The distributions of patients between treatment arms and sites are summarized in Table 1. Treatment allocation was well-distributed between groups at all sites.

Thirty-five subjects were lost to follow-up. These were fairly evenly divided between interferon and placebo groups: 24 subjects in the interferon groups and 11 subjects in the placebo group were lost to follow-up.

Table 1: Subject Enrollment by Site

	Placebo	Betaseron		Total
	(N=308)	8 MIU Fixed (N=317)	5 MIU/m² (N=314)	(N=939)
U of New Mexico, NM	15	14	14	43
Washington Univ School of Medicine, MO	14	14	13	41
Yale, CT	14	13	13	40
Stony Brook, Stony Brook, NY	13	14	13	40
Kaiser Permanente, San Diego, CA	13	13	13	39
U British Columbia, Vancouver, Canada	12	13	12	37
U of South Florida, FL	10	12	12	34
Hopital Notre Dame, Montreal, Canada	10	12	11	33
UCSF-Mt Zion, CA	10	10	11	31
U of Chicago, IL	10	11	10	31
UC Irvine, CA	10	10	10	30
Tulane University Medical Center, LA	10	10	10	30
Duke University Medical Center, NC	10	10	10	30
Vanderbilt University, TN	10	10	10	30
University Hospital, London, Ontario, Canada	10	10	10	30
U of Maryland, MD	9	10	10	29
Ohio State University, OH	10	10	9	29
Medical College Wisconsin, WI	9	10	9	28
Henry Ford Hospital, Detroit, MI	8	9	10	27
Oregon Health Sciences University, OR	8	9	8	25
Bowman Gray School of Medicine, NC	8	8	8	24
Dallas Neurological Clinic, TX	8	8	8	24
The Neurology Center, Rockville, MD	8	7	8	23
U of Arizona, AZ	8	7	7	22
U of Alabama, AL	7	7	7	21
Shepherd MS, Atlanta, GA	6	7	7	20
Mayo Clinic, MN	6	7	7	20
Montreal Neurological Inst., Montreal, Canada	6	6	6	18
Lahey Clinic, Burlington, MA	5	6	6	17
UCLA, CA	5	6	5	16
Carolinas Medical Center, NC	5	6	5	16
Maimonides Med Ctr, Brooklyn, NY	6	5	5	16
Thomas Jefferson University, PA	5	5	6	16
Blake Neurology, Englewood, CO	5	4	6	15
Bernard Gimbel MS, Teaneck, NJ	5	4	5	14

Protocol Deviations

Protocol deviations were reported in 34% of placebo subjects and ~40% of Betaseron subjects and are summarized in Table 2. The excesses of violations in the Betaseron groups were driven primarily by factors related to incomplete receipt of study agent (i.e., agent not taken for ≥6 months, interruptions >30 days, < 80% of doses taken in a 6-month interval). Though there appears to be a disparity between groups with respect to non-protocol steroid use, the rate for the Betaseron groups combined is virtually the same as in the placebo group.

Table 2: Protocol Deviations (per sponsor)

	Placebo	Betaseron	
	N (%)	8 MIU N (%)	5 MIU/m^2 N (%)
Eligibility criteria not met	8 (2.6)	8 (2.5)	14 (4.5)
Drug not taken for ≥ 6 months	17 (5.5)	24 (7.6)	22 (7.0)
Interruption of study agent for > 30 days	31 (10.1)	37 (11.7)	49 (15.6)
< 80% of doses taken during a 6-month interval	42 (13.6)	58 (18.3)	66 (21.0)
Use of steroids (non-protocol)	30 (9.7)	17 (5.4)	41 (13.1)
Use of prohibited medications	9 (2.9)	14 (4.4)	13 (4.1)
Missing > 1 scheduled EDSS evaluation per year	27 (8.8)	18 (5.7)	18 (5.7)
Any of the above protocol deviations	105 (34.1)	114 (40.0)	129 (41.1)

Randomization

Adherence to the randomization scheme (block size = 6) was maintained such that balance of treatment assignment within sites was excellent.

Time-On-Study

Time-on-study (enrollment to last date eligible to report exacerbations or adverse events) represents the duration of opportunity for exacerbations or adverse events to occur. For the

complete data set, time-on-study ranged from 1 to 43 months, with a median of 35.6 months. Time-on-study was similar in all treatment groups. Seventy-two percent of subjects completed 33 months or more on study.

Time-At-Risk (CBER Analysis)

Validation of progression at the penultimate study visit required confirmation by a subsequent examination, at least 70 days later, but prior to the final study visit. Thus, the final opportunity for progression was the latest study visit for which an additional visit occurred at least 70 days later. In light of this requirement, CBER analyzed the sponsor's data sets to determine, for each subject, the final opportunity to record a verifiable progression. Median times at-risk were 32.7 months for both the placebo and fixed-dose Betaseron groups, and 32.8 months for the BSA-adjusted Betaseron group.

In the Betaseron combined groups, as well as in the placebo group, 1.9% of subjects had no opportunity for a confirmed progression. There was no prospective rule for such subjects; however, CBER's examination of the sponsor's data sets indicates that these subjects were censored at the time of their final visit (as were subjects lost to follow-up). Although the lack of a prospective plan was an oversight, these 18 subjects, equally divided between treatment groups, were unlikely to importantly affect the study results.

Treatment Discontinuation

Overall, 24% of subjects discontinued the study agent prematurely, and discontinuations were reported at virtually the same frequency in all 3 groups. However, the reported reasons for treatment discontinuation varied between the groups:

More Betaseron subjects than placebo subjects discontinued study agent prematurely due to AEs: 4% of subjects in the placebo group, 9% in the fixed-dose Betaseron group, and 10% in the BSA-adjusted Betaseron group. Adverse events reported as leading to treatment discontinuation in >2 Betaseron subjects included: asthenia, flu syndrome, myasthenia, depression, and injection site reaction. Of note, more subjects in the placebo group discontinued study agent because of progression of disease.

Table 3: Reasons for Treatment Discontinuation

	Placebo	Betaseron	
	(N=308)	8 MIU Fixed (N=317)	5 MIU/m² (N=314)
Adverse event	12 (4%)	30 (9%)	32 (10%)
Asthenia	1 (0.3%)	6 (1.9%)	5 (1.6%)
Flu syndrome	0 (0%)	2 (0.6%)	3 (1.0%)
Myasthenia	0 (0%)	3 (0.9%)	2 (0.6%)
Depression	1 (0.3%)	2 (0.6%)	1 (0.3%)
Injection site reaction	0 (0%)	2 (0.6%)	2 (0.6%)
other AE	10 (3.2%)	15 (4.7%)	19 (6.1%)
Death	1 (<1%)	4 (1%)	1 (<1%)
Other	4 (1%)	5 (2%)	4 (1%)
Progression of disease	26 (8%)	15 (5%)	13 (4%)
Protocol deviation	4 (1%)	3 (1%)	2 (1%)
Withdrawal of consent, non-compliance, lost to follow-up	28 (9%)	22 (7%)	23 (7%)
Completed study	233 (76%)	238 (75%)	239 (76%)

Reviewer's Comment(s): These data are very similar to those of the EU Study, in which twice as many Betaseron subjects as placebo subjects prematurely discontinued study agent due to AEs, and twice as many placebo subjects discontinued the study agent because of progressive disease.

Treatments Administered

Overall, 75% of subjects completed the study on drug, and 95% of scheduled doses were administered. These rates were consistent for the 3 treatment groups.

CBER explored the datasets with the intent of comparing the doses received by the 2 active treatment groups on a BSA-adjusted basis. Calculation of the dose received by each subject was not straightforward, however, because subjects increased their dose of study agent over the initial 4 weeks on study, and because the doses of study agent administered after the "ramp-up" period were not consistent in all subjects. Using the dataset "dosel.xpt," CBER defined the "dose" of each subject as the predominant dose administered at the end of the study.

The interferon doses received by subjects in the 8 MIU and 5 MIU/m^2 groups were similar on a BSA-adjusted basis, though on the whole, the fixed-dose group received slightly less interferon on a BSA-adjusted basis. Specifically, for the 5 MIU/m^2 group, the median dose (25^{th}, 75^{th} percentile) was 5.0 (4.9, 5.1) MIU/m^2. For the 8 MIU group, the median dose (25^{th}, 75^{th} percentile) was 4.4 (4.0, 4.8) MIU/m^2. This is a small difference that is probably not meaningful. By design, subjects in the BSA-adjusted group were to receive 5 MIU/m^2, and, as expected, 95% of these subjects received doses within 20% of 5 MIU/m^2. For the fixed dosing group, however, 76% of subjects received a dose within 20% of 5.0 MIU/m^2

Blinding

Blinding questionnaires were completed by treating physicians, evaluating physicians and subjects at Week 12 and at study termination. Respondents were asked to guess treatment assignment; however, a "don't know" selection was available to discourage random guessing.

Data are available from ~95% of subjects at Week 12, 90% of subjects at study termination, and 98% of treating and evaluation physicians at both time points. At end-of-study, 26% of patients, 45% of treating physicians and 85% of evaluating physicians guessed "don't know." The high proportion of evaluating physicians providing a "don't know" response, together with the fact that this response was maintained between Week 12 and end-of-study, suggest that the blind was preserved in the critical group.

Subjects tended to provide responses of "don't know" more frequently in the placebo group (31% at end-of study) than the Betaseron groups (24% at end-of-study).

Reviewer's Comment:
A more conservative analysis considers that there is potential bias favoring a response of "don't know," because a correct guess suggests unblinding and potential bias on the part of study participants. CBER re-analyzed the data, combining the two Betaseron groups and eliminating "don't know" responses from the analysis (Table 4). Evaluating physicians, treatment physicians and subjects guessed correctly 53-67%, 79-90% and 46-91% of the time, respectively. It is noteworthy that at Week 12, subjects who received Betaseron and their treating physicians were able to guess the correct treatment assignment >90% of the time. Thus, these data suggest substantial unblinding for treating physicians and for subjects who are assigned to Betaseron, but maintenance of the blind for the critical evaluation physicians.

Study Population: Baseline Characteristics

Baseline demographic characteristics and MS disease status are summarized by treatment group in Table 5. In general, the characteristics are typical of the SPMS patient population. Age and race were well-balanced between treatment groups. There was a slight excess of females in the fixed-dose Betaseron group. MS tends to follow a more benign course in women; therefore, there is the potential for a favorable effect in the Betaseron group.

The groups were generally balanced with respect to baseline disease status. Duration of MS and SPMS were very slightly longer, and T2 lesion area was slightly higher in the placebo group, suggesting the possibility of more advanced baseline disease in the placebo group. These factors would tend to bias the results in favor of the interferon group. There was good balance in terms of pre-study exacerbation rates.

Table 4: CBER Analysis of Blinding Questionnaire - "Don't Know" Responses Eliminated and Betaseron Groups Combined

			Actual Treatment	
			Placebo	Betaseron
Evaluating Physician				
WEEK 12	Guessed Tx	**Placebo**	**66.7%**	45.9%
		Betaseron	33.3%	**54.1%**
WEEK 156	Guessed Tx	**Placebo**	**65.1%**	47.3%
		Betaseron	34.9%	**52.7%**
Treating Physician				
WEEK 12	Guessed Tx	**Placebo**	**79.2%**	9.7%
		Betaseron	20.8%	**90.3%**
WEEK 156	Guessed Tx	**Placebo**	**79.3%**	20.0%
		Betaseron	20.7%	**80.0%**
Subject				
WEEK 12	Guessed Tx	**Placebo**	**45.9%**	8.9%
		Betaseron	54.1%	**91.1%**
WEEK 156	Guessed Tx	**Placebo**	**63.2%**	18.1%
		Betaseron	36.8%	**81.9%**

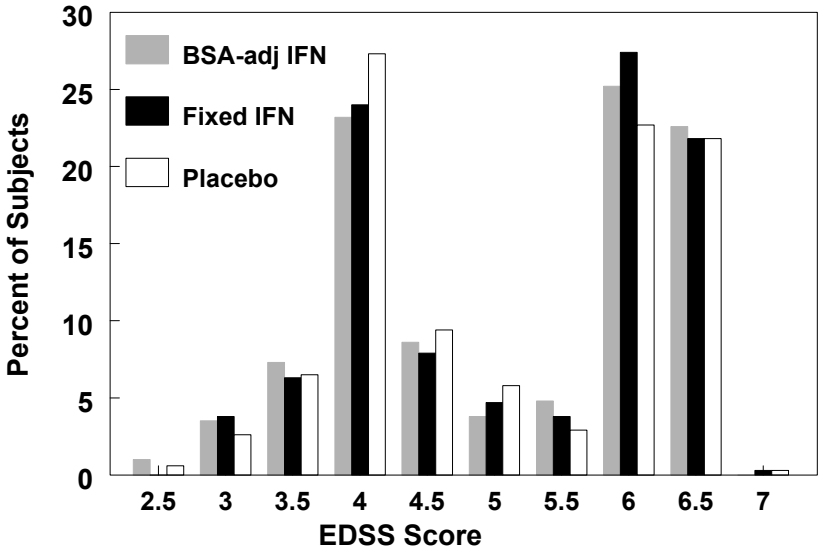
Pre-study EDSS constitutes a key baseline parameter, because it is indicative of the level of pre-existing disability, which is related to progression of disability (the primary efficacy endpoint). There was good balance in the 3 treatment groups with respect to baseline EDSS.

Table 5: Demographic Characteristics and Baseline Disease Status

Characteristic		Placebo	Betaseron	
			8 MIU Fixed	5 MIU/m²
N		308	317	314
Age (years)	Mean (SD)	47.6 (8.2)	46.1 (8.0)	46.8 (8.3)
	Range	24 - 65	22 - 65	28 - 65
% Female		60	66	61
Race				
Asian	N (%)	2 (1)	1 (<1)	0 (0)
Black	N (%)	19 (6)	22 (7)	21 (7)
Caucasian	N (%)	279 (91)	286 (90)	287 (91)
Hispanic	N (%)	6 (2)	3 (1)	6 (2)
Other	N (%)	2 (1)	5 (2)	0 (0)
Duration of MS [years]	Mean (SD)	14.9 (8.3)	14.6 (7.8)	14.5 (8.7)
Duration of SPMS [years]	Mean (SD)	4.1 (3.5)	4.0 (3.3)	4.0 (3.5)
Relapses in prior 2 years	Mean (SD)	0.8 (1.2)	0.8 (1.1)	0.9 (1.6)
0 relapses	N (%)	174 (56)	170 (54)	173 (55)
1 relapse	N (%)	72 (23)	83 (26)	82 (26)
> 2 relapses	N (%)	62 (20)	64 (20)	59 (19)
Baseline EDSS	Mean (SD)	5.1 (1.1)	5.2 (1.1)	5.1 (1.2)
EDSS 2 years prior to entry	Mean (SD)	3.4 (1.4)	3.5 (1.5)	3.4 (1.4)
Δ EDSS during 2 years prior to entry	Mean (SD)	+ 1.7 (0.9)	+ 1.7 (0.9)	+ 1.7 (0.9)
Baseline Scripps score	Mean (SD)	65.1 (11.7)	65.3 (11.7)	64.2 (11.3)
MRI: T2 lesion area [mm^2]	Mean (SD)	4327 (4129)	3805 (3891)	3857 (3967)

CBER plotted individual baseline EDSS scores by treatment group (Figure 1). Overall, the distribution of baseline EDSS scores was not normally distributed across the scale range, but bimodal. The larger peak was at the 6.0 to 6.5 range, with a secondary peak (encompassing about half as many subjects) at EDSS = 4.0. Epidemiological studies indicate that the duration at grade EDSS 6.0 is relatively long in comparison with grades 4.0 and 5.0, hence a greater frequency at EDSS = 6.0 is expected. With respect to balance between treatment arms, there tended to a slight excess of placebo subjects with baseline EDSS scores of 4.0 to 5.0, inclusive, and slightly fewer placebo subjects with EDSS ≥ 5.5 (Figure 1). These are small imbalances, that are unlikely to importantly influence the results of the study.

Figure 1: Baseline EDSS by Treatment Group



In summary, with respect to demographic parameters and baseline disease status, there are minor imbalances in gender, duration of disease and T2 volume favoring the interferon groups, and a small imbalance in baseline EDSS favoring the placebo group. Given the small magnitude of these imbalances, however, their effect, if any, would be expected to be very limited.

Comparison of baseline characteristics with those of the EU Study (ME 93079)

Compared with the EU study (ME 93079), subjects in the NA trial were ~5 years older, with an MS duration that was ~3 years longer. This difference was largely due to time spent in the secondary progressive phase of MS. Whereas the EU subjects were reported to be in the secondary phase of MS for a mean of 1.3 years, the mean duration of SPMS was 4.0 years for subjects in the NA study. The distribution of % of subjects with durations ranging from <1 year to > 4 years is shown in Table 6. This is a key difference, in that the subject population of the EU study had entered more recently into the secondary progressive phase of MS. Baseline EDSS in the NA study was ~0.5 point lower than in the EU study.

Table 6: Comparison of duration of SPMS with Study ME 93079

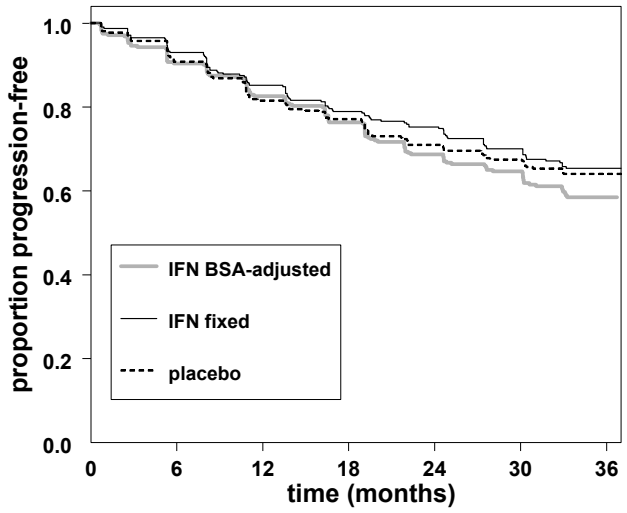
	% of subjects	
	EU Study	NA Study
< 1 year	22%	9%
1 to < 2 years	35%	22%
2 to < 3 years	15%	22%
3 to < 4 years	10%	12%
≥ 4 years	18%	35%

Results:

Primary efficacy endpoint:

In total, there were 333 progressions in the 939 subjects (an overall progression rate of 35%). There were 106 progressions (34%) recorded in the placebo group, 124 (39%) in the BSA-adjusted Betaseron group, and 103 (32%) in the fixed-dose Betaseron group. The Kaplan-Meier analysis is shown in Figure 2. For the logrank test, the prospectively-defined method of analysis, the results were not statistically significant: all Betaseron vs. placebo: p = 0.71; fixed-dose Betaseron vs. placebo: p = 0.61; BSA-adjusted Betaseron vs. placebo: p = 0.26.

Figure 2: Primary Efficacy Endpoint – Time to Disease Progression



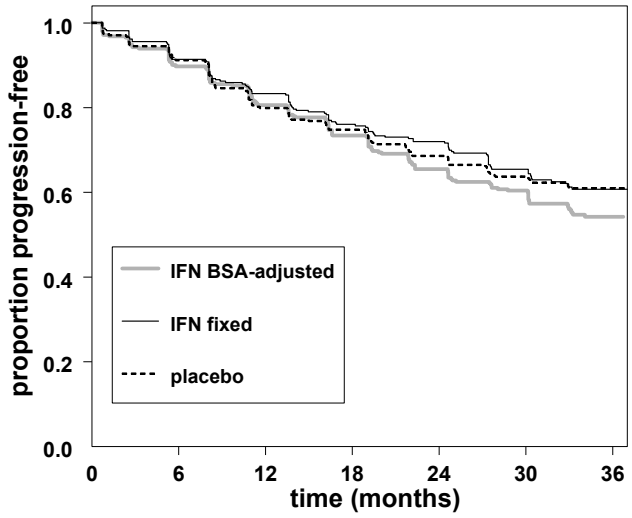
Supportive analyses, including Cox regression stratified for site, and Cox regression stratified for site and controlled for baseline EDSS and duration of MS were likewise not statistically significant. Analyses for the efficacy evaluable dataset were not statistically significant.

Exploratory analyses on the primary endpoint:

1. Alternate definition of progression

In study ME 93079, progression required confirmation of the increase in EDSS over a minimum of 3 months instead of 6 months, and EDSS scores obtained within the first 90 days of an ongoing relapse were not counted. Using the ME 93079 definition of progression, there were slightly greater numbers of progressions, but there remained no difference among treatment groups (Figure 3).

Figure 3: Time to Progression – Using European Study's Definition of Progression



2. Analysis of primary endpoint by center – CBER Analysis

CBER combined the two interferon treatment groups together, and assessed the proportions of subjects with progression and the relative risk (RR) of progression by study site. The RR (interferon/placebo) ranged from a minimum of 0.32 at Stony Brook, to a maximum of 4.50 at U.C. Irvine. The degree of variability between sites was not unexpected for a study of this size and nature. No particular patterns emerged with respect to the RR of progression in terms of site size, site geographic location, university centers versus private centers, etc. Overall, the analysis demonstrated fairly symmetric variability about a mean RR of 1.05 (95% CI 0.87 – 1.26).

3. Potential effect of pre-study exacerbations on apparent progression – CBER Analysis

The inclusion of subjects with RRMS have the potential to confound the study. Subjects in whom there were no exacerbations reported in the 2 years prior to study may represent a "purer" population of SPMS subjects. As such, in an exploratory analysis, CBER combined the two interferon treatment groups and assessed time to progression in the subset of subjects with no reported exacerbations in the 2-year pre-screening period. For these subjects, representing slightly more than half the total study population (placebo n=174; interferon n=343), there was a trend in favor of *placebo* in delaying time to progression. Specifically, 33% of placebo-treated subjects experienced progression at 3 years, compared to 41% of interferon-treated subjects. Conversely, for the subjects *with* reported exacerbations in the 2-year period prior to screening (placebo n=134; interferon n=288), subjects in the interferon groups tended to have a *delayed* time to progression (rates of progression over 3 years were 34% and 39% for the interferon and placebo groups, respectively).

4. Potential effects of concomitant medications

Numerous medications have been used for symptomatic benefit in MS. Although the use of these agents has not been shown to affect the natural course of the disease, concomitant medication use has to potential to alter symptoms, perception of symptoms, and/or physical findings in MS. Thus, concomitant medication use, if unbalanced in the study, could confound the results by differentially affecting the EDSS. Agents with the potential to importantly influence patient assessments are summarized below:

- Limb spasticity, ataxia and weakness are amenable to pharmacologic modification with muscle relaxants, antispasmodics, and benzodiazepines, and such modification could be sufficiently effective to alter the EDSS score of a subject.

- Fatigue is a frequent symptom which may alter mobility in MS patients, thereby affecting EDSS. Dopaminergic agents (typically amantadine) and psychostimulants (typically pemoline) have been used to reduce fatigue in this disease.

- Bladder and bowel function are components of Kurtzke Functional Systems (FS) assessments, which are, in turn, components of EDSS assessments. Therefore, use of parasympathomimetic and antiadrenergic agents are potentially confounding variables.

- Numerous agents prescribed for neuropsychologic impairment and pain management, including benzodiazepines, antidepressants and anticonvulsants, also can affect FS assessments and EDSS.

- Prohibited immune modifiers such as glatiramer acetate, interferon β-1a and interferon β-1b could confound the results of the study if their frequency of use was significant and unbalanced.

The sponsor found statistically significant imbalances in the use of androgens, antipruritics, astringents, tranquilizers and "blocking agents." The latter category included interferon beta-1a (Avonex),----------and cyclophosphamide. The numbers of subjects who received cyclophosphamide were small, and included 5 subjects in each of the interferon groups, and 2 subjects in the placebo group. (Of note, of the 5 subjects in each of the interferon groups who received cyclophosphamide, 1 received the agent for beast cancer.)

Non-study interferon use – CBER Analysis
Non-study interferon use was reported in 21, 3 and 6 subjects in the placebo, fixed interferon and BSA-adjusted interferon groups, respectively. The 21 placebo subjects with reported non-study interferon use represent 6.8% of the total placebo subjects.

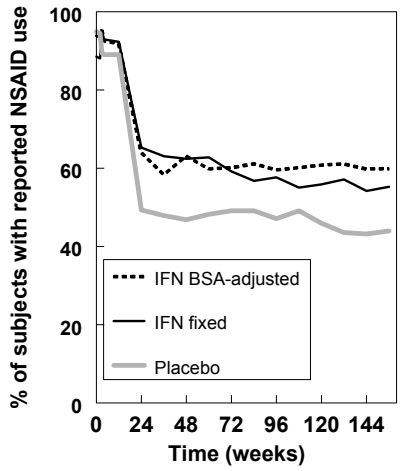
CBER performed an exploratory analysis on the primary disability endpoint, omitting the 21 subjects in the placebo group with reported non-protocol interferon use. With the elimination of these subjects from the placebo group, there was a shift towards *less* disease progression. Thus, the rate of progression in placebo subjects with non-protocol interferon use tended to be disproportionately high. Thus, the study disability results were not confounded by non-protocol interferon use in placebo subjects.

Non-steroidal anti-inflammatory drugs (NSAIDs) – CBER Analysis

The sponsor tabulated concomitant medication use by individual medication and by drug class, as percent of subjects who reported use while on study. For these analyses, therefore, subjects with a single reported use of a concomitant medication were given the same weight as subjects with reported use throughout the study.

CBER reassessed reports of concomitant medication use by weeks on study, and found an imbalance in the use of NSAIDs, with greater use in the interferon groups after Week 12 (Figure 4). For this analysis, aspirin and aspirin-containing products were considered to be NSAIDs, and all products were considered to have equal weight.

Figure 4: Percent of subjects with reported NSAID use by study week



In light of this imbalance, time to progression was assessed in subgroups of NSAID use and treatment assignment in an exploratory analysis. For each subject, the number of visits during which NSAID use was reported was divided by the total number of visits during which information on concomitant medication use was collected, and expressed as a fraction. For the study overall, the median fraction of visits during which NSAID use was reported was 0.706, corresponding to approximately 25 months of NSAID use. A dichotomous analysis was performed for which subjects were divided into NSAID "users" and "non-users" on the basis of whether NSAID use was reported in less than or greater than (or equal to) the median fraction of visits. Additional analyses were performed on the basis of NSAID use tertiles. There was a slight trend in favor of less progression of disability in association with NSAID use. When NSAID use was subdivided into tertiles, the trend was apparent only for the highest tertile of NSAID use. The analyses show no convincing effect of NSAIDs on progression, and suggest that the primary efficacy measure was not confounded by differential rates of NSAID use in the interferon and placebo groups.

5. Subgroups defined by time of entry into study – CBER Analysis

CBER performed exploratory analyses on the primary efficacy endpoint with respect to time of study entry. Subjects were divided into two groups based on study entry before or after the median enrollment date (August 5, 1996). The two interferon groups were combined. There was a trend towards an interferon treatment effect in the subjects enrolled between August 5, 1996 and April 1, 1997 (second "half" of study enrollment). As expected, this was counterbalanced by an opposite trend in subjects enrolled between August 29, 1995 and August 5, 1996 (first "half"). These differences were minimal, however, and likely reflect the increased variance of subgroup analyses with decreased sample sizes.

6. Subgroups in U.S. vs. Canada – CBER Analysis

To assess the possibility of differential efficacy by country, CBER assessed time to progression at US sites (821 subjects) and Canadian sites (188 subjects) separately. The results in both U.S. and Canadian sites were similar to the those of the study as a whole.

7. Reconciliation of the discordant results of the NA and EU studies – Exploratory analyses to identify demographic and disease-specific characteristics predictive of a Betaseron treatment response

The hallmark of SPMS is slowly progressive disability - - disability that can have a greater day-to-day impact on patients than relapses (the latter become progressively less frequent in the

later stages of MS). The apparent absence of a salutary Betaseron treatment effect in this study stands in contrast to the results of the EU study, wherein there was a demonstrable treatment effect. CBER's Complete Review letter of June 4, 1999 stressed the importance of reconciling these discordant study results.

Comparison of the NA and EU studies suggested disparities in baseline demographic and disease-specific characteristics between the patient populations. The recognition of particular characteristics, associated with a Betaseron treatment response and consistent across both studies, could potentially permit the identification of patient sub-population(s) more likely to receive treatment benefit. The sponsor pointed out issues that might explain disparities between the 2 patient populations:

- Patients in the EU study were enrolled at a time when there were no approved disease-modifying therapies for any form of MS in the EU. The NA study was initiated 1 year later, and prior to its recruitment, both interferon beta-1a and glatiramer acetate had been approved for RRMS. The sponsor suggests that the approval of therapies for RRMS may have led to some off-label use in SPMS, predominantly for patients with more disease activity. Thus, a negative selection would leave the pool of SPMS patients with overall less disease activity.
- There was further reduction in the pool of SPMS patients because of competition for recruitment with clinical investigations of other MS therapies.

With reduction in the size of the potential pool of subjects – patients naïve to interferons, glatamer, and investigational therapies – the sponsor speculates that the NA study enrolled patients destined to manifest relatively low levels of disease activity, as measured by relapse rate and progression of functional impairment.

Consistent with this premise, comparison of the placebo arms of the 2 studies suggests that NA subjects had had more limited MRI activity and less frequent relapses than subjects in the EU study. The sponsor hypothesizes that the inability to demonstrate a treatment effect on progression of disability in the NA study was related, in part, to the tendency for these patients to have "burnt out" disease, relative to patients in the EU study.

The sponsor performed a multivariate discriminant analysis on numerous baseline variables, designed to ascertain whether the NA and EU study subjects differed with respect to baseline characteristics, and to provide an estimate of the difference. The baseline variables that were found to significantly contribute to the ability of the model to separate subjects into "NA-like" versus "EU-like" groups included: age, baseline EDSS, BSA, gender, Δ EDSS over the 2 years prior to study, lymphocyte count, and relapses within the 2 years prior to study entry (the latter to the 0.5 power). Each variable was associated with a coefficient, allowing the calculation of a summed score for each subject. Subjects with scores higher than a specific cut-off were classified as "NA-like;" subjects with lower scores were classified as "EU-like."

Based on this categorization, the proportions of subjects with confirmed progression are shown in Table 7, top, displayed by study (EU and NA), subject classification (i.e., "EU-like" versus "NA-like"), and treatment. The sponsor elected to show data for only the subjects in the NA study who received fixed doses of Betaseron and placebo.

Based on an exploratory analysis of doses administered (see page 27), CBER takes the view that there is little intrinsic difference between subjects who received fixed doses of

Table 7: Proportion of Subjects with Confirmed Progression – Subjects Categorized as "NA-Like" or "EU-Like"

		EU Study	
Sponsor's Analysis:		**"NA-like"**	**"EU-like"**
group	**Betaseron**	36/103 (35.0%)	93/257 (36.2%)
	placebo	43/109 (39.4%)	120/249 (48.2%)
		NA Study (fixed doses only)	
		"NA-like"	**"EU-like"**
group	**Betaseron**	62/210 (29.5%)	41/107 (38.3%)
	placebo	32/114 (28.1%)	17/40 (42.5%)
CBER Analysis:			
		NA Study (fixed and weight-adjusted doses)	
		"NA-like"	**"EU-like"**
group	**Betaseron**	139/414 (33.6%)	88/217 (40.6%)
	placebo	73/224 (32.6%)	33/84 (39.3%)

Betaseron/placebo versus subjects who received weight-adjusted doses. Thus, CBER extended the analysis to include all subjects in the NA study, including both fixed-dose and BSA-adjusted dose subjects (Table 7, bottom).

Approximately 70% of subjects in the EU study were categorized as "EU-like," with the remainder classified as "NA-like." The Betaseron treatment effect was observed primarily in "EU-like" subjects, with only a slight positive trend in "NA-like" subjects.

In the NA study, approximately 70% of subjects were categorized as "NA-like," and there is no trend suggesting a Betaseron treatment effect in these subjects. For the 30% of subjects classified as "EU-like," the sponsor's tabulation suggests a slight trend in favor of Betaseron in preventing progression of disability. CBER's analyses of the combined study population (fixed and BSA-adjusted dosing, Table 7, bottom) shows no apparent effect of Betaseron on prevention of disability progression in "EU-like" subjects in the NA study (i.e., 40.6% and 39.3% of "EU-like" subjects with progression in the Betaseron and placebo groups, respectively).

For the combined studies, the sponsor performed a Cox proportional hazards regression analysis, attempting to identify subgroups of patients, based on single or multiple characteristics, with a higher probability of receiving benefit from Betaseron. The baseline variables considered in the model included: duration of MS, duration of SPMS, relapses in the 2 years prior to study, gender, Δ EDSS over the 2 years prior to study, baseline MRI lesion volume, baseline lymphocyte count, BSA, age, baseline EDSS score, and baseline gadolinium enhancing lesions.

The baseline characteristics that appeared to identify subgroups with consistent treatment effects for the 2 studies included: relapses in the 2 years prior to study, Δ EDSS over the 2 years prior to study, and gadolinium enhancing lesions at baseline. The sponsor construed these factors as being consistent with "recent disease activity." Table 8 provides a summary of

the RRs for these variables, for both the NA and EU studies, with RR denoting relative risk of progression with Betaseron relative to placebo.

Table 8: Relative Risk (RR) of progression for various subgroups (Betaseron relative to placebo)

variable	value of variable	NA Study		EU Study	
		N	RR	N	RR
Prior Relapses	No	517	1.28	212	0.80
	Yes	422	0.81	506	0.67
Prior Δ EDSS	< 1	317	1.63	212	0.90
	≥ 1	622	0.83	221	0.55
Enhancing Lesions	No	98	1.68	60	0.85
	Yes	65	0.59	65	0.78

For the EU study, Betaseron-treated patients with any of these factors (relapses or increases in EDSS in the 2 years prior to study; gadolinium enhancing lesions at baseline) exhibited decreased RR of progression. For the complementary groups (i.e., no relapses or change in EDSS score in the 2 years prior to study, etc.), the RRs favored Betaseron, though less persuasively.

Consistent with the EU study results, the NA study data suggested decreased RR of disability progression in Betaseron-treated subjects with recent disease activity, manifested as relapses or increased EDSS score within 2 years, or enhancing lesions at baseline. However, in contrast to the EU study, the RR in complementary subgroups (i.e., no recent disease activity) were consistent with *increased* risk of disability progression.

These post-hoc analyses suggest an association between recent disease activity and treatment benefit, a hypothesis that may be worthy of additional testing. In light of the divergent results of the NA and EU studies with respect to progression of disability, the magnitude of benefit, if any, can not be quantified. No clearly identifiable demographic or disease-specific factors were predictive of a consistent benefit of Betaseron with respect to disability outcome.

8. Exploratory analyses to characterize a possible relation between early MRI responses and benefit of Betaseron on disability progression

As part of CBER's complete response letter of January 2001, CBER requested exploratory analyses to assess the potential relation between MRI responses to Betaseron early after treatment initiation, and progression of disability with continued treatment. If such analyses suggested an association, a consideration for further study would be a design featuring population enrichment based on retention of patients in whom early MRI findings suggested Betaseron responsiveness.

The sponsor assessed the relation between gadolinium enhancing lesions during the initial 6 months of treatment and progression of disability over 3 years. These analyses were limited to the frequent MRI subgroups of the 2 studies (n=163 NA study; n=125 EU study). Using two different definitions of an early MRI response, assessment of subgroups defined on the basis of this response, the sponsor found no appreciable differences in terms of eventual progression of disability.

Reviewer's Comments:

The sponsor's first definition defined "worsened" MRI activity as any increase in the monthly rate (months 1-6) over baseline, with other subjects characterized as "flat or improved." This definition did not address patients with early improvement in MRI lesions, per se.

The sponsor's second definition was based on comparisons between numbers of lesions at baseline and during months 1-6. An early MRI response was defined as "worsened" for a ≥ 0.2 lesions/month increase from baseline; "flat" as within ± 0.2 lesions/month of baseline; and "improved" for a ≥ 0.2 lesions/month decrease from baseline. The sponsor displayed the data by study, and divided Betaseron-treated NA subjects by group (fixed versus BSA-adjusted dose). As such, the N's were very limited in each subgroup. CBER combined all Betaseron-treated subjects in the NA study, based on the presumption that the 2 dosing regimens are essentially equivalent. In addition, CBER considers the salient comparison to be improved versus not improved. Thus, the "worse" and "flat" subgroups were combined into a single "not improved" subgroup, and compared to the "improved" subgroup (Table 9).

Table 9: Relation Between Early MRI Activity Status and Progression of Disability

	NA		EU		NA & EU	
MRI Activity Status	**Placebo**	**Betaseron**	**Placebo**	**Betaseron**	**Placebo**	**Betaseron**
Improved	9/13 (69%)	13/38 (34%)	8/18 (44%)	15/30 (50%)	17/31 (55%)	28/68 (41%)
Not improved	11/42 (26%)	30/70 (43%)	24/43 (56%)	15/34 (44%)	35/85 (41%)	45/104 (43%)

For the NA study, Betaseron-treated patients who had shown an early MRI response ("improved") experienced a 34% rate of disability progression, whereas the rate of progression was 43% for subjects lacking early evidence of MRI improvement. This trend was reversed in the EU study, with rates of progression of 50% and 44% in subjects with and without evidence of early MRI improvement, respectively. Thus, for Betaseron-treated subjects across both studies combined, rates of progression were essentially the same, irrespective of early improvements in MRI parameters. Thus, an early response as assessed by MRI parameters does not appear to predict Betaseron responsiveness with respect to subsequent progression of disability.

The data for the placebo subjects are relevant to the potential relation between early MRI findings and subsequent disability. The data show opposite trends for the NA and EU studies, and overall do not suggest a correlation between MRI findings and subsequent disability progression.

Overall, the results of the NA study do not support the conclusion that Betaseron slows progression of disability in SPMS. Multiple patient subsets based on demographic factors, baseline characteristics, and early post-treatment response indicators were examined. These included gender, age, disease duration, indicators of disease activity prior to study enrollment, and MRI measures both at baseline and early changes from baseline following treatment. No factors were identified that were predictive of a subgroup of patients showing decreased progression of disability in response to Betaseron. The study does not support the conclusion that Betaseron treatment delays progression of disability in any identifiable population or subpopulation of patients.

Secondary Endpoints

1. Change in mean EDSS score – baseline to endpoint

No significant differences were seen among treatment groups in the analyses of change in mean EDSS scores from baseline to end-study. The values for change in mean scores were 0.62, 0.53, and 0.72 in the placebo, fixed- and BSA-adjusted Betaseron groups, respectively. ANOVA analyses yielded p-values ranging from 0.21 to 0.63 for the comparisons of each active treatment group to placebo, and for all Betaseron to placebo. Similarly, the GEE analysis did not reveal a Betaseron treatment effect.

2. Annual relapse rate

The mean annual relapse rates for the placebo, fixed- and BSA-adjusted dose groups were 0.28, 0.16, and 0.20, respectively. Overall, Betaseron-treated patients experienced a lower annual relapse rate than placebo patients (p=0.014 for all Betaseron vs. placebo). The difference between the fixed-dose and placebo groups was statistically significant (p=0.009), whereas the difference between the BSA-adjusted Betaseron and placebo groups was not. The effects of baseline relapse rate, baseline EDSS, gender and age at onset of MS were explored using an ANCOVA model. After adjusting for these baseline characteristics, the pairwise comparisons of active treatment to placebo were significant at p<0.04.

CBER exploratory analyses on the relapse endpoint

Analyses of the complete data set were performed by CBER using raw data in -------ransport datasets. A total of 580 relapses were reported with a valid start date: 159 in the fixed-dose interferon group, 179 in the BSA-adjusted interferon group, and 242 in control subjects. The annual relapse rate was calculated as the number of relapses divided by time-on-study (in days) multiplied by 365.25 days per year (Table 10). Annual exacerbation rates were also assessed by exacerbation severity (Table 11).

Table 10: Annual Relapse Rates

	Placebo	Betaseron		p-values		
		8 MIU	5 MIU/m^2			
	(N=308)	(N=317)	(N=314)			
Total number of relapses						
	231	145	171			
Annual relapse rate						
Median	0	0	0	0.0143[1]	0.0034[2]	all Betaseron vs placebo
Mean	0.28	0.16	0.20	0.0091	0.0035	8 MIU vs placebo
SD	0.51	0.34	0.39	0.1088	0.0356	5 MIU/m^2 vs placebo
Range	0.00 - 3.94	0.00 - 3.03	0.00 - 3.84			
Number of subjects with relapses						
	116 (38%)	91 (29%)	105 (33%)	0.0444[3]	all Betaseron vs placebo	
				0.018	8 MIU vs placebo	
				0.2951	5 MIU/m^2 vs placebo	
Number of relapses per subject						
0	192 (62%)	226 (71%)	209 (67%)	0.0009[4]	all Betaseron vs placebo	
1	56 (18%)	59 (19%)	64 (20%)	0.0013	8 MIU vs placebo	
2	36 (12%)	20 (6%)	25 (8%)	0.0281	5 MIU/m^2 vs placebo	
3	9 (3%)	7 (2%)	10 (3%)			
4	10 (3%)	4 (1%)	4 (1%)			
5+	5 (2%)	1 (<1%)	2 (1%)			

Only investigator-verified relapses are used.

[1] ANOVA model based on square-root of annual relapse rate including site, treatment, and treatment-by-study site interaction

[2] ANCOVA - same model adjusted for baseline relapse rate, baseline EDSS, gender, and age at onset of MS

[3] CMH test stratified for site

[4] CMH ANOVA test stratified for site

Table 11: CBER Analysis of Annual Relapse Rates by Relapse Severity

Relapse Severity *	Placebo mean ± SD	Placebo median	Betaseron mean ± SD	Betaseron median	Relative Rate **
mild	0.09 ± 0.24	0.00	0.07 ± 0.19	0.00	0.72
moderate	0.14 ± 0.31	0.00	0.10 ± 0.23	0.00	0.67
severe	0.05 ± 0.22	0.00	0.03 ± 0.18	0.00	0.56
moderate or severe	0.20 ± 0.42	0.00	0.13 ± 0.31	0.00	0.64
all	0.29 ± 0.54	0.00	0.19 ± 0.38	0.00	0.66

* includes all relapses (verified and non-verified)

** annual rate in Betaseron groups/annual rate in placebo group

The effect of Betaseron on annual relapse rate is consistent across the three severity categories, with decreases in the mean annual exacerbation rate in the range of 28% (mild exacerbations) to 44% (severe exacerbations).

Relapse rates in the NA study were ~40% of those reported in Study ME 93079. The interferon treatment effect, however, was very consistent between the two studies.

An analysis of severity of worst relapse by subject showed, for the combined interferon groups, a net shift of 3.4% of subjects from both the "severe" and "moderate" categories to the "none" and "mild" categories (Table 12).

Table 12: Severity of Worst Relapse

Severity of Worst Relapse	Placebo n=308	Betaseron (both) n=631
mild	27 (8.8%)	59 (9.4%)
moderate	64 (20.8%)	110 (17.4%)
severe	27 (8.8%)	34 (5.4%)
unknown	1 (0.3%)	3 (0.5%)
none	189 (61.4%)	425 (67.4%)

3. MRI - Absolute change in T2 lesion area

T2 lesions are thought to represent fixed disease due to prior attacks, and the T2 volume represents an approximation of the total extent of these lesions. An assessment of the correlation between lesion volume and activity and clinical benefit was a phase 4 commitment of the sponsor.

Table 13: Change in MRI T2 Area

	Placebo (N=308)	Betaseron 8 MIU Fixed (N=317)	Betaseron 5 MIU/m² (N=314)	p-value	
Baseline					
N	308	313	312	0.07[1]	all Betaseron vs placebo
Median (cm^2)	32.8	24.7	25.1	0.03	8 MIU vs placebo
Mean (cm^2)	43.3	38.1	38.6	0.32	5 MIU/m^2 vs placebo
Change from baseline to week 48					
N	168	161	167	<0.001[2]	all Betaseron vs placebo
Median (cm^2)	1.7	-0.2	0.1		
Mean (cm^2)	3.6	-0.5	-0.1		
Change from baseline to week 96					
N	160	157	172	<0.001[2]	all Betaseron vs placebo
Median (cm^2)	2.8	0	0		
Mean (cm^2)	6.3	0.2	0.5		
Change from baseline to week 156					
N	248	248	261	<0.001[2]	all Betaseron vs placebo
Median (cm^2)	2.1	0.1	0.1	<0.001	8 MIU vs placebo
Mean (cm^2)	6.4	1.3	1.4	<0.001	5 MIU/m^2 vs placebo
Change from baseline to last on-study scan					
N	290	294	301	<0.001[2]	all Betaseron vs placebo
Median (cm^2)	2.3	0	0.1	<0.001	8 MIU vs placebo
Mean (cm^2)	6.4	1.1	1.3	<0.001	5 MIU/m^2 vs placebo

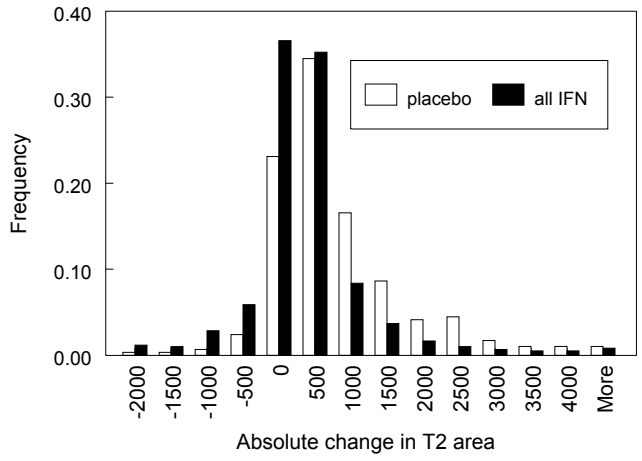
[1] ANOVA model based on ranked data including treatment, site, and treatment by site

[2] ANCOVA model based on ranked data including treatment, site, and treatment by site, adjusted for baseline lesion area

The sponsor provided interim data with this amendment to the BLA supplement, summarized in Table 13. Treatment groups were to be compared for absolute and percent change in T2 lesion area from baseline to last scan using an analysis of covariance (ANCOVA) model based on ranked data and adjusted for the ranks of the baseline lesion burden. The main effects in the model were treatment, study site, and the treatment-by-study site interaction. After examination of the data and performing ANCOVAs based on both actual and ranked data, the sponsor determined that the model based on the ranked data was the most appropriate. A *t*-test was used to test within-group change from baseline.

The protocol called for a supportive repeated measures analysis on lesion area over time. However, Week 48 and Week 96 scans for approximately half of the subjects had not been quantified as of the submission of this amendment. Thus, the repeated measures analysis was not performed.

Figure 5: Change in T2 Area by Treatment Group



At baseline, T2 lesion area tended to be greater in subjects assigned to placebo than in subjects assigned to Betaseron. Based on an ANCOVA model that included ranked data for treatment, site, treatment by site and baseline lesion area, there is a strong treatment effect for both Betaseron groups. Whereas there was, on average, an increase in lesion area of 16% in placebo-treated subjects, increases in mean lesion area in Betaseron-treated subjects were in the range of 5 – 6%.

Exploratory analyses on the T2 volume endpoint

CBER assessed the distributions of changes in MRI T2 area (Figure 5). For all categories representing a change in lesion area of 500 mm^2 or less (including all categories representing *decreases* in lesion area), there are greater fractions of interferon-treated subjects. Conversely, for all categories representing increases in T2 lesion area of 1000 mm^2 and above, there are excesses in placebo-treated subjects.

4. Annual newly active lesion rate

To assess the effect of interferon on new lesion formation, a 163-subject frequent MRI subgroup underwent exams before and after administration of a gadolinium contrast agent every 4 weeks throughout the study. Newly active lesions were defined as unique lesions that were identified one or more times but which were not in the previous scan. The annualized newly active lesion rate was calculated by computing the cumulative number of newly active lesions and dividing by the number of scans used in the cumulative computation. This resulted in the average number of newly active lesions per scan which was then multiplied by 13 (the number of expected scans per year) to calculate the annual activity rate. Baseline activity was computed as the average of the newly active lesions at screen and Day 1. No correction was made for missing values.

Annual newly active lesion rates were calculated for 55 subjects in the placebo group, and 54 subjects in each of the interferon groups. The mean annual newly active lesion rates were 18.7,

6.4, and 4.5 in the placebo, fixed-dose, and BSA-adjusted dose interferon groups, respectively. The sponsor's p-values for the comparisons of all Betaseron, fixed-dose Betaseron, and BSA-adjusted Betaseron vs. placebo were 0.0001, 0.0030, and 0.0001, respectively. These findings were supported by a GEE analysis.

5. Change in composite neuropsychologic score

Results for change in composite neuropsychologic score showed no differences between treatment groups at any time point. The mean change from baseline to endpoint for the placebo, fixed-dose, and BSA-adjusted dose interferon groups was -0.32, -0.28, and -0.30, respectively, from baseline means of -0.83, -0.80, and -0.81. The patterns of gradual decrease over time appeared to be similar between the treatment groups.

Tertiary Endpoints

1. Tertiary EDSS Endpoints

Proportion of subjects with confirmed disease progression

Similar proportions of subjects experienced progressions in the placebo, fixed-dose, and BSA-adjusted interferon groups (34%, 32%, and 39%, respectively.) Supportive analyses also showed similar results, with no clear pattern for analyses incorporating stratification for site, or adjustment or stratification for baseline EDSS. No statistically significant differences were seen using either the primary method (CMH test stratified for site) or a supportive method (CMH test stratified for baseline EDSS).

The sponsor's exploratory analyses included:

- Progression was defined as per the definition used in study ME 93079. The proportions of subjects with 3-month confirmed progression were 37% in the placebo and fixed-dose interferon groups, and 44% in the BSA-adjusted interferon group.
- Active treatment groups were compared to their respective placebo matching subgroups, i.e., fixed placebo vs. fixed interferon, BSA-adjusted placebo vs. BSA-adjusted interferon. The proportion of subjects in the placebo fixed-dose group who progressed was identical to that in the active fixed dose group (32%). Similarly, the proportion of subjects in the placebo BSA-adjusted dose group who progressed was nearly equal to that in the BSA-adjusted interferon group (37% vs. 39%, respectively).
- The data for 6-month confirmed progression were analyzed using CMH tests stratified for pre-study and on-study relapsing status. The proportion of subjects who relapsed in the 2 years prior to the study who experienced 6-month progression was 37%, 30%, and 35% in the placebo, fixed-dose, and BSA-adjusted interferon groups, respectively. No significant differences were seen in the pairwise comparisons of all Betaseron, fixed-dose and BSA-adjusted interferon to placebo.

2. Tertiary Relapse Endpoints

Time to first relapse

The time to experiencing the first on-study relapse was significantly prolonged in Betaseron-treated subjects (*p*=0.023 for all Betaseron vs. placebo). The difference between the fixed-dose and placebo groups was also statistically significant (*p*=0.010), whereas the difference between the BSA-adjusted interferon and placebo groups was not (*p*=0.195).

Distribution of relapse severities

The distribution of Scripps-determined relapse severities (categorized as none, mild, moderate or severe) showed statistically significant differences between all Betaseron-treated subjects and placebo (p=0.021) and fixed-dose interferon vs. placebo group (p=0.006), but not for the BSA-adjusted interferon group vs. placebo (p=0.240). The distribution of relapses was less severe in the all Betaseron and fixed-dose groups compared to the placebo group. No statistically significant differences between treatment groups were seen in the comparisons of investigator-determined relapse severities.

Proportion of subjects with moderate or severe relapses

Overall, only 45 relapses were moderate or severe according to the Scripps criteria. The proportions of subjects in the placebo, fixed-dose, and BSA-adjusted interferon groups who experienced moderate or severe relapses (based on Scripps-determined severity) were 6%, 3% and 4%, respectively. These differences were not statistically significant. For investigator-determined relapses, the proportion of subjects in the placebo, fixed-dose, and BSA-adjusted interferon groups with moderate or severe relapses was 30%, 21%, and 24%, respectively.

Annual rate of moderate or severe relapses

The mean annual rate of Scripps-determined moderate or severe relapses was 0.02, 0.01, and 0.02 in the placebo, fixed-dose, and BSA-adjusted interferon groups, respectively. The mean annual rate of investigator-determined moderate or severe relapses was substantially higher - - 0.19, 0.10, and 0.14 in the placebo, fixed-dose, and BSA-adjusted interferon groups, respectively.

Reviewer's Comment(s): Rates of investigator-determined moderate or severe relapses in EU Study ME 93079 were 0.31 in the interferon group and 0.49 in the placebo group. Thus, the rates observed in the NA study were roughly 40% of those in the EU study.

Mean duration of relapses per subject

The mean duration of relapses per subject was 29.1, 20.3, and 28.5 days in the placebo, fixed-dose, and BSA-adjusted interferon groups, respectively. These differences were not statistically significant.

Total number of days per subject spent in relapses

The mean total number of days per subject spent in relapses was 55.4, 33.5, and 45.8 days in the placebo, fixed-dose, and BSA-adjusted interferon groups, respectively. The comparison of all interferon subjects to placebo subjects was not statistically significant.

Reviewer's Comment(s): Of note, the mean duration of relapses and numbers of days spent in relapses was roughly half those observed in Study ME 93079

Number of days per subject spent in a moderate or severe relapse

The mean total number of days spent in Scripps-determined moderate or severe relapses were 4.0, 1.5 and 4.2 in the placebo, fixed-dose, and BSA-adjusted interferon groups, respectively. The analogous values based on investigator determinations were 37.1, 22.0, and 32.6 days, respectively. These differences were not statistically significant.

Number of steroid courses per subject

There were 1.10, 0.64, and 0.97 steroid courses per subjects in the placebo, fixed-dose, and BSA-adjusted interferon groups, respectively. The difference in the number of steroid courses per subject between the fixed-dose interferon group and the placebo group was statistically significant (p=0.001), as was the difference between all Betaseron and placebo (p=0.018). The comparison of the BSA-adjusted interferon group to placebo was not statistically significant (p=0.41).

Proportion of subjects with systemic corticosteroid use

The proportions of subjects receiving systemic corticosteroids were 37% for the fixed-dose interferon group, and 46% for both the BSA-adjusted interferon and placebo groups. The comparison of all interferon-treated subjects to placebo subjects was not statistically significant.

MS-related hospitalizations

Two hundred fifty-two subjects (27% of the total study population) were hospitalized during the study, with approximately (123 subjects, 13% of the total study population) for MS-related reasons. There were no significant differences among treatment groups in the number of MS-related hospitalizations per subject.

Proportion of subjects with time in the ICU

Only 3% of subjects overall experienced time in the ICU during the study, and no interferon treatment effect was observed for this endpoint.

3. Tertiary MRI endpoints

Data for these endpoints were obtained from the frequent MRI cohort of 163 subjects. Two of the planned endpoints (change in annual non-enhancing MRI lesion area and recurrent lesion rate) were not assessed because the procedures could not be performed by the UBC group.

Per subject proportion of scans with specified lesion activities

The mean per subject proportion of scans with active lesions in the placebo, fixed-dose, and BSA-adjusted interferon groups was 33.7, 17.5, and 15.4, respectively. The mean per subject proportion of scans with enhancing lesions (either new or persistent) was 36.0, 19.2, and 16.4, respectively. Both parameters showed high levels of statistical significance for the pairwise comparisons of all Betaseron, fixed-dose, and BSA-adjusted interferon to placebo.

Proportion of subjects with any active scan

The proportion of subjects with any active scan (a scan showing a newly active lesion) reflects the likelihood of having $\geq$1 newly active lesion during the course of the study, and were 69%, 70% and 56% for the placebo, fixed-dose, and BSA-adjusted interferon groups, respectively. These differences were not statistically significant.

Annual rates for specified types of lesions

The endpoints with respect to rates for specified types of lesions showed statistically significant treatment effects for both doses of Betaseron and for all Betaseron vs. placebo. These endpoints included annual rates for new, newly enlarging, newly enhancing, persistently enlarging, persistently enhancing, and persistently active lesions.

Correlations between clinical, MRI and other outcomes

The sponsor computed Spearman rank correlation coefficients (SRCC) to explore relationships between clinical, MRI and other outcome measures, and in particular to evaluate MRI as a surrogate marker of disease activity in MS. The SRCC for the primary correlations were not strong. The SRCC for the change in EDSS score at endpoint vs. change in MRI lesion area at endpoint was only 0.15. The SRCC for the correlation between annual relapse rate and annual rate of enhancing lesions was 0.27. Supportive analyses for the correlations between annual relapse rate with 3 alternate MRI outcomes (proportion of enhancing scans, change in lesion area at endpoint, and newly enhancing lesion rate) ranged from 0.15 to 0.29.

Reviewer's Comment(s): These data do not demonstrate a good correlation between MRI findings and clinical outcomes, and do not support the concept that MRI findings can be used as surrogate markers of MS disease activity.

Change for individual instruments of the neuropsychologic assessment

The change from baseline was analyzed for the individual instruments of the neuropsychologic assessment. No significant differences were seen in the comparisons of placebo to Betaseron-treated subjects for any of these analyses, including the analyses using the normal control data.

Quality of life

There were no significant differences between treatment groups at any time point for the SF-36 and SIP dimensions, measures of health-related QOL. A slight decrease was observed in QOL over time in all groups; however, the changes were small and clinically insignificant. Consistent with previous studies, subjects had a considerably lower QOL in the SF-36 physical domain than in the mental domain. Specifically, the baseline means for the SF-36 physical component score were lower than the respective mean in the US general population by 2 standard deviations. Baseline means of all 3 treatment groups for the SF-36 *mental score* were similar to that of the US general population, with no significant inter-group differences.

Change in Environmental Status Scale

The ESS was administered at yearly intervals as a measure of the impact of MS on social roles and employment. The mean change from baseline to endpoint was small for all treatment groups, with no significant differences between groups.

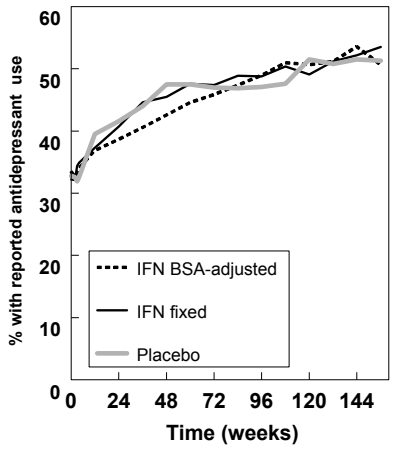
New use of antidepressants

The proportions of subjects newly treated with antidepressants were similar in the 3 treatment groups: 29%, 29%, and 31%, for the placebo, fixed-dose, and BSA-adjusted interferon groups, respectively.

Because of the association of interferons (particularly alpha interferons) and depression, CBER analyzed the sponsor's data files for % of subjects with reported antidepressant use by reporting period. In addition to anti-depressants, the CBER analysis included St. John's Wort.

There was a gradual increase in antidepressant use in all 3 treatment groups, from ~33% at the beginning of the study to ~50% at end-of-study, with no apparent differences between treatment groups (Figure 6).

Figure 6: Antidepressant Use Vs. Time



Global impression of improvement, worsening, and stability

A global MS status survey was administered at Week 12 and Week 156 to obtain, from subjects, treating physicians and evaluating physicians, a global impression of improvement, worsening and stability. Neither the subject's nor the treating physician's impressions showed differences among treatment groups at either time point. Though the evaluating physician's assessment at Week 12 did show statistically significant differences in favor of Betaseron, this apparent treatment effect did not persist at the Week 156 assessment. Moreover, the multiplicity of analyses undermines the strength of this observation.

Clinical pharmacology

Subjects from selected sites were enrolled in a pharmacokinetic substudy. At present, pharmacokinetic data cannot be obtained due to the lack of an assay capable of measuring interferon beta-1b at the levels present in subject samples.

Safety

Extent of exposure

All 939 patients comprising the ITT population received at least one dose of study agent. Thus, the ITT and safety populations are identical. At the time the study was terminated, the mean time on study and time on agent were ~1000 and ~900 days, respectively, and similar for all treatment groups (Table 14). Dose reductions were more frequent in the Betaseron groups.

Table 14: Extent of exposure to study agent

	Placebo	Betaseron	
		8 MIU Fixed	5 MIU/m²
N	308	317	314
Days on study (mean)	1022	998	1013
Days on test agent (mean)	922	895	918
% of doses taken (mean)	96%	95%	94%
% subjects with dose reductions	5%	15%	22%
assigned dose (MIU)	-	8.0	9.6
administered dose (MIU)	-	6.6	7.8

Deaths

Seven deaths occurred during the study period, including two subjects in the placebo group and five subjects in the Betaseron groups.

Placebo group: A 55 year-old female died of "arteriosclerosis" on week 11, and a 55 year-old male died of pneumonia, approximately 2.5 years into the study. CIOMS forms were not provided for these 2 placebo deaths.

Fixed-dose Betaseron group: Two female subjects, one age 43 and another age 44, were reported to die of "cardiac arrest;" a 36 year-old female died of a pulmonary embolism, and a 58 year-old male committed suicide:

- A 43 year-old female had a history of serious gastrointestinal problems, pre-dating entry into the study, and had a number of serious co-morbidities at the time of her sudden death. Six months prior to her demise, she experienced mesenteric ischemia and underwent small bowel resection, with numerous complications. At the time of her death, she had a chronically infected abdominal wound site, fever, anemia, increasing azotemia, and an uncharacterized seizure disorder.

- A 44 year-old female who had been receiving Betaseron for 30 months, had sudden death without apparent acute antecedent illness. There was no autopsy, and the cause of death is unknown.

- A 36 year-old wheelchair-bound female who had received Betaseron for 5 weeks experienced sudden death due to autopsy-proven deep venous thrombosis and pulmonary embolism.

- A 58 year-old male committed suicide via a self-inflicted gunshot wound to the head. There was a significant family history of depression; the subject's father had committed suicide by the same means.

BSA-adjusted Betaseron group: A 56 year-old male died due to metastatic lung cancer 23 weeks into the study. No CIOMS form was provided. A 36 year-old male died with a presumptive diagnosis of pulmonary embolism:

- A 36 year-old male with obesity and diabetes developed supraventricular tachycardia, hypotension and hypoxemia. A presumptive diagnosis of pulmonary embolism was made on the basis of echocardiographic evidence of acute right atrial and right ventricular dilatation shortly before death.

Serious Adverse Events (SAEs)

The sponsor reported subject incidences of SAEs of 28% in the placebo group, 27% in the fixed-dose Betaseron group, and 28% in the BSA-adjusted Betaseron group. None of the SAEs reported during the study in 2% or more of the total study population was significantly associated with Betaseron treatment.

Table 15: Serious adverse events with reported incidence ≥ 1% in Betaseron-treated subjects – CBER analysis

	Placebo (N=308)	Betaseron (N=631)
asthenia/myasthenia	10.9%	11.7%
fever	3.5%	3.6%
urinary tract infection	3.3%	2.6%
hypertonia	2.7%	2.6%
abnormal gait	2.5%	2.3%
incoordination	2.5%	1.6%
abdominal pain	1.9%	0.7%
* pneumonia	1.9%	2.6%
pathological fracture	1.7%	1.0%
* ataxia	1.7%	2.3%
pain	1.6%	1.6%
paresthesia	1.3%	0.7%
dyspnea	1.3%	1.0%
back pain	1.1%	0.7%
* dehydration	1.1%	2.3%
convulsion	1.1%	0.7%

Reviewer's Comment: In exploring that SAS datasets provided, CBER could not verify all of the sponsor's numbers for SAEs. Moreover, there was considerable overlap between "asthenia," which was reported under "body as a whole" and "myasthenia," which was reported under "musculoskeletal system." Thus, CBER combined these categories in its analyses of the------dataset ADVRSL.xpt.

Serious AEs reported in ≥ 1% of Betaseron-treated subjects are shown in Table 15. Both Betaseron groups are combined (n = 631). The most common SAEs were asthenia/myasthenia, fever, urinary tract infection and hypertonia. Serious AEs for which the frequency in the Betaseron groups exceeded that in the placebo group are denoted with an asterisk, and include pneumonia, ataxia and dehydration. Given the multiplicity of these comparisons and the lack of known potential underlying pharmacologic mechanisms, these differences should probably be ascribed to chance. Thus, there do not appear to be important differences in the incidences of SAEs between the Betaseron groups and the placebo group.

Severe AEs

The sponsor did not provide analyses of severe AEs. CBER analyzed severe AEs from the ------dataset ADVRSL.xpt, combining the two Betaseron groups, and combining the categories asthenia and myasthenia, because of overlap, as noted for SAEs, above. Table 16 lists, in descending order of frequency, all reported severe AEs with a combined incidence in the Betaseron groups of ≥ 1%. Severe AEs with a reported frequency in the Betaseron groups exceeding the reported frequency in the placebo group by 25% are denoted with an asterisk (*). Severe AEs in this category include headache, leukopenia, urinary tract infection, abdominal pain, fever, injection site reaction, chills, rhinitis, tremor, convulsion, anxiety, somnolence and urinary frequency. An additional severe AE of concern (because of its serious nature and a trend toward increased frequency in Betaseron-treated subjects) is depression. Rates of severe depression are 7.9% and 6.8% in the Betaseron and placebo groups, respectively.

All Adverse Events

CBER's analysis of AEs is shown in Table 17. AEs with a reported frequency in the combined Betaseron groups ≥ 15% are shown on the left; AEs with reported frequency < 15% but with a frequency in the Betaseron groups exceeding that of the control group by ≥ 25% are shown at right.

Notable adverse events associated with Betaseron use in the present study can be divided into two categories, based on a comparison with adverse events noted in the Warnings and/or Adverse Reactions sections of the existing labeling for Betaseron for RRMS:

1. Notable AEs identified in the present study and also noted in the Warnings and/or Adverse Reactions sections of current Betaseron RRMS labeling. These include: abdominal pain, anxiety, chills, confusion, convulsions, depression, dyspnea, fever, headache, injection site hypersensitivity/ inflammation/ necrosis/ pain/ reaction, lymphadenopathy, malaise, myalgia, palpitations, peripheral vascular disorders, somnolence, suicide, and menstrual disorders.

2. Notable AEs identified in the present study but NOT noted in the Warnings and/or Adverse Reactions sections of current Betaseron RRMS labeling are tabulated in Table 18. Frequencies in boldface font represent AEs listed because of frequencies of **severe** events.

With respect to possible labeling changes, taking into consideration the multiplicity of these comparisons, their relative rates, and their medical significance, the simple inclusion of events with higher frequencies in the adverse event table would probably be adequate, without the addition of any specific warnings or precautions.

Table 16: Severe AEs –CBER Analysis

	Betaseron (N=631)	Placebo (N=308)
asthenia/myasthenia	39.3%	35.4%
hypertonia	12.5%	11.4%
* headache	9.8%	7.1%
pain	8.6%	7.8%
abnormal gait	8.1%	9.4%
depression	7.9%	6.8%
* leukopenia	5.4%	1.3%
ataxia	4.8%	9.1%
incoordination	4.8%	4.9%
back pain	4.0%	6.2%
paresthesia	3.5%	6.2%
* urinary tract infection	3.5%	2.0%
dizziness	3.3%	3.3%
* abdominal pain	2.7%	1.3%
arthralgia	2.7%	2.6%
* fever	2.5%	2.0%
urinary incontinence	2.5%	2.6%
pathological fracture	2.4%	2.3%
* injection site reaction	2.2%	0.3%
nausea	2.2%	4.9%
* chills	2.1%	0.3%
diarrhea	2.1%	2.0%
chest pain	1.7%	2.0%
migraine	1.7%	2.0%
dyspnea	1.6%	1.3%
flu syndrome	1.6%	2.6%
myalgia	1.6%	1.3%
constipation	1.4%	2.0%
dehydration	1.4%	2.3%
fecal incontinence	1.4%	1.3%
* rhinitis	1.4%	1.0%
* tremor	1.4%	1.0%
* convulsion	1.3%	0.3%
pneumonia	1.3%	2.3%
* anxiety	1.1%	0.7%
diplopia	1.1%	1.0%
* somnolence	1.1%	0.3%
* urinary frequency	1.1%	0.3%
urinary retention	1.1%	2.0%
vomiting	1.1%	3.3%

* Frequency in Betaseron groups exceeds that in the placebo group by > 25%

Table 17: Subject incidence of adverse events – CBER analysis

AEs: reported frequency > 15% in Betaseron group

	Betaseron (N=631)	placebo (N=308)
pain	65.0%	61.4%
asthenia	63.1%	60.1%
hypertonia	61.7%	60.1%
myasthenia	58.5%	61.4%
headache	57.5%	47.4%
* injection site reaction	53.6%	14.9%
* injection site inflammation	51.5%	7.5%
* flu syndrome	45.2%	34.7%
depression	44.4%	44.2%
paresthesia	43.1%	45.1%
pharyngitis	40.9%	39.9%
arthralgia	38.4%	38.0%
rhinitis	34.2%	43.8%
incoordination	34.1%	32.8%
urinary tract infection	33.8%	34.1%
back pain	33.3%	32.5%
nausea	32.0%	31.5%
insomnia	31.1%	26.3%
abnormal gait	30.7%	32.8%
fever	29.6%	24.7%
dizziness	28.5%	27.0%
accidental injury	27.4%	26.6%
* myalgia	26.9%	18.5%
* rash	26.6%	21.1%
* leukopenia	25.2%	8.4%
ataxia	24.7%	33.4%
constipation	24.4%	25.3%
peripheral edema	23.0%	20.1%
diarrhea	22.8%	19.5%
urinary incontinence	22.4%	21.1%
* chills	22.0%	11.7%
* abdominal pain	21.7%	17.2%
skin disorder	20.0%	18.8%
urinary urgency	19.8%	17.5%
* injection site pain	19.5%	9.1%
dyspepsia	19.2%	18.5%
sinusitis	17.9%	18.5%
ecchymosis	16.3%	24.4%
amblyopia	15.5%	15.3%
cough increased	15.1%	15.6%

AEs: reported frequency < 15% in Betaseron group, but frequency in Betaseron groups exceeds frequency in the placebo group by > 25%

	Betaseron (N=631)	placebo (N=308)
* chest pain	14.7%	8.1%
* bronchitis	10.1%	7.5%
* dyspnea	9.7%	6.8%
* lymphadenopathy	9.4%	5.2%
* lung disorder	7.1%	5.5%
* thinking abnormal	7.0%	4.2%
* leg cramps	6.7%	2.9%
* peripheral vascular disorder	6.7%	3.6%
* malaise	6.5%	2.0%
* menstrual disorder	6.3%	4.9%
* vaginitis	6.0%	3.6%
* mouth ulceration	5.6%	3.9%
* injection site necrosis	5.2%	0.3%
* alopecia	4.9%	3.6%
* dry mouth	4.8%	2.3%
* fungal dermatitis	4.8%	2.3%
* injection site hypersensitivity	4.8%	1.6%
* palpitation	4.8%	2.3%
* tinnitus	4.8%	3.3%
* dry skin	4.6%	3.6%
* epistaxis	4.6%	2.6%
* neuralgia	4.6%	2.9%
* tachycardia	4.1%	2.9%
* eye disorder	3.7%	2.3%
* anemia	3.5%	2.3%
* injection site edema	3.2%	0.7%
* syncope	3.2%	2.3%
* eczema	3.0%	1.6%
* liver function tests abnormal	3.0%	1.0%
* tendon disorder	3.0%	1.6%
* amenorrhea	2.9%	1.0%
* skin ulcer	2.9%	2.0%
* arthritis	2.5%	1.3%
* bone disorder	2.4%	1.6%
* moniliasis	2.4%	1.6%
* movement disorder	2.4%	1.3%
* skin discoloration	2.4%	0.3%
* aphthous stomatitis	2.1%	1.0%
* colitis	2.1%	1.3%
* paralysis	2.1%	1.3%
* SGPT increased	2.1%	0.3%
* urinary tract disorder	2.1%	1.6%

* Frequency in Betaseron groups exceeds that in the placebo group by ≥ 25%

Specific Abnormalities

Flu-like symptoms

The flu-like symptom complex was defined as the presence of flu syndrome or at least two of the following events concurrently: fever, chills, myalgia or sweating. Of the subjects in the placebo group, 37% reported flu-like symptom complex at some point during the study, whereas 51% of Betaseron-treated subjects reported flu-like symptom complex at some time. Over the course of the study, approximately twice as many Betaseron-treated subjects as placebo subjects reported flu-like symptoms (Figure 7).

Injection Site Reactions

The classification of AEs as "injection site reactions" included AEs with the terms "injection site..." "edema," "hemorrhage," "hypersensitivity," "inflammation," "mass," "necrosis," "pain," and "reaction." Eight-eight percent (88%) of Betaseron-treated subjects, and 39% of placebo subjects reported at ≥ 1 injection site reaction during the course of the study. Figure 8 shows the weekly prevalence of injection site reactions for the 3 groups as a function of time (CBER analysis). For Betaseron-treated subjects, the mean number of injection site reactions per subject over the course of the study was 2.2 (median = 1.0). Of the 1359 unique reported injection site reactions in interferon-treated subjects, 32 (2.4%) were classified as severe, 315 (23.2%) were classified as moderate, and 1012 (74.5%) were classed as mild. Four events (0.3%) were SAEs. For placebo subjects, the mean number of injection site reactions was 0.7 (median = 0), with 2 (0.9%) classified as severe events and none classified as SAEs.

Table 18: Notable AEs, not represented in current Betaseron RRMS labeling

Adverse Event	Betaseron (N=631)	placebo (N=308)
rash	26.6%	21.1%
tremor	12.2%	10.7%
tremor (severe)	**1.4%**	**1.0%**
bronchitis	10.1%	7.5%
thinking abnormal	7.0%	4.2%
leg cramps	6.7%	2.9%
vaginitis	6.0%	3.6%
dry mouth	4.8%	2.3%
fungal dermatitis	4.8%	2.3%
urinary tract infection	33.8%	34.1%
urinary tract infection (severe)	**3.5%**	**2.0%**
injection site edema	3.2%	0.7%
skin discoloration	2.4%	0.3%
urinary frequency	13.0%	12.3%
urinary frequency (severe)	**1.1%**	**0.3%**

Figure 7: Flu-like symptom complex vs. time – CBER Analysis

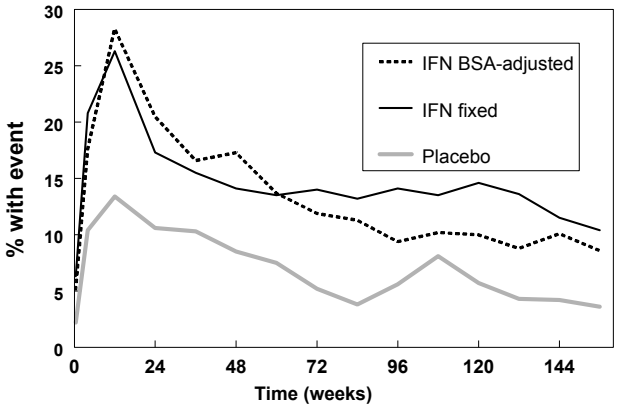
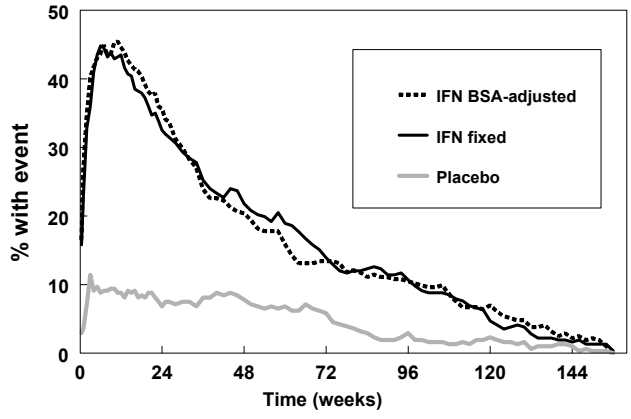


Figure 8: Injection site reactions vs. time – CBER Analysis



Depression

Depression is not uncommon in MS patients, particularly in a patient population with more advanced disease (as in this study). Any therapy that can exacerbate this propensity has the potential to convey significant harm to these patients. Depression and suicidal tendencies are reported in the labeling as adverse reactions for Betaseron, identified in the Warning section. Three Betaseron-treated subjects attempted suicide, of which one was completed. There were no suicide attempts reported in the placebo group.

Depression was reported as an AE (serious and non-serious) in 126 subjects (41%) in the placebo group and 42% of subjects in the Betaseron groups. Depression as an SAE was reported in1% of subjects in the placebo group and 0.6% of subjects in the Betaseron groups. Psychotic depression was reported as an AE in two subjects in each of the three groups (1% per group).

The mean BDI scores for all three groups fell within the "minimal" category (score of 0-9) throughout most of the study. Only rarely did a mean value increase to "mild depression" (score of 10-16). Overall, mean BDI scores showed an initial decrease and remained relatively stable thereafter until the last visit when all three groups showed a slight increase. There were no statistically significant differences between either treatment group and the placebo group in the mean BDI score change from baseline to any time point. There was also no difference among the three groups in the number of subjects classified as "at risk for suicide" at any time during the study according to the BDI: placebo, 15%; fixed-dose Betaseron, 17%; and BSA-adjusted Betaseron, 17%. The percentages of subjects in the three groups with moderate or severe depression (score of >16) at any time during the study were also similar across all groups: 38%, 34% and 39% in the placebo, fixed-dose and BSA-adjusted Betaseron groups.

The sponsor found no difference among the three groups with respect to the proportion of subjects with antidepressant use at baseline (32-35%) or newly treated with antidepressants during the study (29-31%). (A subject must have used the antidepressant for $\geq$ 14 days to be considered as having been treated with an antidepressant).

CBER performed an exploratory analysis to assess the proportions of subjects with reported antidepressant use by week, and included reported use of St. John's Wort (an herbal OTC agent, not uncommonly used to treat depression/mood disorders). The CBER analysis (see Figure 6, page 46) is substantially in agreement with the sponsor's conclusions.

Menstrual Disorders

Menstrual disorders were reported at similar frequencies in the Betaseron and placebo groups. Metrorrhagia was reported in 9% of Betaseron-treated subjects and 10% of placebo subjects. Menorrhagia was reported in 8% and 7% of subjects in the Betaseron and placebo groups, respectively. Dysmenorrhea was reported in 6% and 5% of subjects in the Betaseron and placebo groups, respectively. "Other" menstrual disorders were reported in 10% of subjects in the Betaseron treatment groups, and 8% of subjects in the placebo group. Overall, therefore, there was no apparent association between Betaseron and menstrual disorders in the NA study.

Thyroid Disorders

At baseline, approximately 5% of subjects in each treatment group reported thyroid disease. During the course of the study, hypothyroidism was reported in 1.9% and 1.3% of subjects in the Betaseron and placebo groups, respectively. Thyroid neoplasia was reported in 3 (0.5%) Betaseron-treated subjects and none of the controls. Thyroiditis was reported in 1 Betaseron-

treated subject (0.2%) and none of the controls. Hyperthyroidism was reported in 1 Betaseron-treated subject and none of the control subjects. Thus, there appeared to be possible associations between Betaseron treatment and thyroid neoplasia, thyroiditis, and hyperthyroidism; however, all were fairly rare events, and no definite conclusions can be drawn.

Thromboembolic Disorders

There were 2 deaths related to pulmonary embolism in the Betaseron groups, and no reports in the placebo group. Deep venous thrombosis was reported in 1 Betaseron subject (0.2%) and 2 placebo subjects (0.6%). Thrombophlebitis was reported in 2 Betaseron subjects (0.3%) and none of the control subjects. Arterial thrombosis was not reported in Betaseron subjects, but reported in 1 placebo subject. A number of these subjects were at risk of thromboembolic phenomena, because of immobility and obesity.

Reviewer's Comment(s): Subjects 243 and 485 in the EU study were hospitalized with pulmonary embolism. Thus, over both studies, there were 4 reports of pulmonary embolism in 991 Betaseron-treated subjects (0.4%), versus none in placebo subjects. Though these findings may seem concerning on the surface, they must be considered in the context of the multiplicity of AE endpoints analyzed, and the overall low frequency of this event.

Laboratory Abnormalities

Elevations in hepatic abnormalities and hematologic abnormalities were reported more frequently in Betaseron-treated subjects:

Serum glutamic oxaloacetic transaminase (SGOT)

Toxicity of Grade 2 or greater was reported in 2.4% and 1.0% of subjects in the Betaseron and placebo groups, respectively. Grade 3 toxicity was reported in <1% of subjects in both groups, and grade 4 toxicity was not reported.

Serum glutamic pyruvic transaminase (SGPT)

Grade 2 or higher toxicity was reported in 4.6% and 1.6% of Betaseron-treated and control groups, respectively. For grade 3 or 4 toxicity, the respective rates were 0.6% and 1.1. Grade 4 toxicity was reported in only 1 subject (Betaseron-treated).

Lymphocytes

Lymphopenia of grade 2 or greater was reported in 53% of Betaseron-treated subjects, and 23% of control subjects. Grade 3 or 4 lymphopenia was reported in 24% of Betaseron-treated subjects and 7% of control subjects, and the corresponding rates of grade 4 toxicity were 6% and 2%. Thus, serious lymphopenia was reported approximately 3 times more frequently in Betaseron-treated subjects than control subjects.

Platelets

Mild thrombocytopenia (platelet count 75 – 90 X 10^9/L) was reported in 100% of subjects in all treatment groups. Platelet counts <75 X 10^9/L were reported in 2 Betaseron-treated subjects (0.3%), and for both of these subjects, thrombocytopenia was severe (<25 X 10^9/L). These were no reports of platelet counts <75 X 10^9/L in the placebo group.

Neutralizing antibodies (NAB)

The sponsor provided data on antibody development at titer cutoffs of 20, 100, and 400 NU/mL, and noted that a titer of 20 NU/mL was the lowest quantifiable titer for the assay. The overall

incidence of NAB in this subject population was lower than that observed in the previous study of relapsing-remitting subjects. At a titer cutoff of 20 NU/mL, 28% of subjects receiving Betaseron were classified as NAB positive at some point in the study, compared to 1 (<1%) subject receiving placebo. Seroconversion occurred relatively uniformly throughout the first two years of the study, reaching a plateau beyond that point. The incidence of NAB positivity was 14% at one year and 27% at 2 years. After 2 years, only an additional 1% of Betaseron-treated subjects converted.

Multiple analyses were performed to assess the interaction between the development of NAB and treatment response measured by clinical and MRI outcome measures, and no consistent effects of NAB positivity were detected.

The sponsor assessed the frequencies of a number of adverse events in Betaseron-treated subjects with respect to NAB positivity. A dichotomous analysis ("never NAB+" versus "ever NAB+") was used, with NAB positivity defined as a titer of $\geq$ 1:20.

NAB positivity was associated with reports of elevated hepatic transaminases as adverse events as follows: For subjects who developed NAB positivity, increased SGOT was reported as an adverse event in 2.9%, versus 0.7% in those who were never NAB+. With respect to elevated SGPT as an adverse event, the respective frequencies were 4.6% and 1.1%. Of note, however, approximately two-thirds of NAB+ subjects who developed elevated transaminases as an adverse event experienced the peak in liver enzymes prior to NAB development.

Conversely, injection site necrosis was less frequent in subjects who were ever NAB+ versus those who were never NAB+ (0.6% versus 7.0%, respectively).

Reviewer's Comments: In light of the multiplicity of adverse events tested, these findings might be easily ascribed to chance; however, the sponsor performed similar analyses for the EU study dataset at CBER's request, and the results were largely consistent with those of the NA study. In the EU study, there were positive associations between NAB positivity and elevated SGOT and SGPT as adverse events, as well as negative associations between NAB positivity and injection site necrosis/reactions. The significance of these observations is unknown, however, and the addition of specific precautions to the package insert with respect to NAB positivity does not appear warranted at this time.

Financial Disclosure Information

One principal investigator and one sub-investigator received direct support from Berlex Laboratories to conduct phase 4 studies. Together, these investigators participated at sites that enrolled 46 subjects. In light of the study design, and given the negative efficacy results, these relationships are not likely to have importantly influenced the study results or data interpretation.

Assessment

Study design

This was a randomized, double-blind, placebo-controlled, multicenter study designed to assess the efficacy and safety of 2 dosing regimens of Betaseron in U.S. and Canadian outpatients with SPMS. Subjects were randomly allocated to receive, on a QOD schedule, treatment with placebo, Betaseron 8 MIU (the licensed dose for RRMS), or Betaseron 5 MIU/m^2 BSA. The placebo group was further randomized so that half of the subjects received a fixed 1.0 mL of placebo solution and half received a BSA-adjusted fraction of 1.0 mL. Doses were to be

escalated incrementally over 3 weeks, and subjects were to receive ibuprofen 1200 mg/day X 7 weeks to reduce flu-like symptoms, and to decrease the potential for patient unblinding. The planned treatment duration was 156 weeks.

The primary endpoint was progression of disability as determined by EDSS scores. There were five secondary endpoints. These were not prospectively designated as to importance, and included change in mean EDSS score, annual relapse rate, absolute change in MRI T2 lesion area, annual rate of newly active MRI lesions, and change in composite neuropsychologic score from baseline to endpoint.

The primary efficacy endpoint and the secondary endpoint on change in EDSS over the course of the study directly address the issue of progression.

The secondary exacerbation endpoint is generally important in MS trials; however, interferon is licensed as an agent to decrease the frequency of exacerbations, and a salutary effect is to be expected based on the experience from prior studies.

MRI studies are both objective and quantitative, and may be performed on a serial basis. MRI-related endpoints have been utilized frequently in clinical studies as surrogates of MS disease burden and activity; however, their clinical meaningfulness remains unproven. Thus, the secondary MRI endpoints are viewed as providing important supportive information.

The tertiary endpoints were numerous, and not prospectively designated with respect to relative importance. Many were designed to support the primary and secondary endpoints. Two tertiary endpoints (Beck Depression Inventory and new use of antidepressants) were designed to assess depression.

Study conduct

The study enrolled 939 subjects and was terminated early on the basis of prospectively planned interim analyses of efficacy and safety at a scheduled IMB meeting in August, 1999. The IMB recommended early study termination, based on the rationale that the primary outcome was unlikely to be changed importantly on the basis of the limited data that had yet to be collected.

Randomization and blinding

Randomization was performed centrally. There were no notable imbalances at any site with respect to the numbers of subjects randomized to treatment group.

The study included measures to maintain the treatment blind by administration of ibuprofen prophylactically to reduce flu-like symptoms associated with interferon use. Compliance with prophylactic ibuprofen early in the study was generally successful, with approximately 90% of subjects in all treatment groups using ibuprofen through the first 12 study weeks. After Week 12, NSAID use decreased considerably. From Week 24 and beyond, NSAID use was fairly consistent as 60% and 50% of subjects in the Betaseron and placebo groups continued NSAIDs, respectively.

Although injection site reactions, flu-like symptoms and the need for continuance of non-steroidal anti-inflammatory drugs could serve to unblind subjects and their treating physicians, specially designated evaluating physicians were used to assess EDSS scores for the primary efficacy endpoint. These physicians were to remain as uninformed as possible with respect to

elements of patient status that might lead to unblinding. Because EDSS criteria are largely independent of patient effort, the primary endpoint is deemed to be largely reliable in this study, even in the presence of patient unblinding. This component of the study design argues for accepting the assessments as unbiased. Analyses of blinding questionnaires suggest substantial unblinding of subjects and treating physicians (privy to patient symptoms, signs and laboratory abnormalities), with maintenance of the blind for evaluating physicians. Because blinding of evaluating physicians is critical for the interpretability of the study, and because evaluating physicians were not able to accurately guess treatment assignment, these results suggest that the EDSS evaluations for the primary efficacy endpoint can be accepted as unbiased.

Study population

The two treatment groups were reasonably well-balanced for demographics and baseline disease status. There were minor imbalances in gender, duration of disease and T2 volume favoring the interferon groups, and a small imbalance in baseline EDSS favoring the placebo group. Given the small magnitude of these imbalances and their opposite direction, their effect, if any, would be expected to be very limited.

The mean EDSS at baseline for all randomized subjects was 5.1 and the mean Scripps score was 64.9, indicating moderate to severe physical disability. The mean proportion of subjects who were relapse-free within the two years prior to the study (55%) and the mean change in EDSS score in the same time period (1.7) were both relatively high, suggesting that the population may have been skewed toward more advanced stages of SPMS. The mean baseline duration of MS (14.7 years), duration of SPMS (4.0 years) and the subject age (46.8 years) are also consistent with more chronic and advanced disease.

Relative to the EU study ME 93079, the subjects in the NA trial were ~5 years older, with an MS duration that was ~3 years longer. This difference was largely due to time spent in the secondary progressive phase of MS. Whereas the EU subjects were reported to be in the secondary phase of MS for a mean of 1.3 years, the mean duration of SPMS was 4.0 years for subjects in the NA study. However, despite having spent more time in the secondary progressive phase of MS, subjects in the NA study had a mean baseline EDSS that was ~0.5 points *lower* than in the EU study. This suggests an inherent difference in the natural history of MS between continents, a difference in the way that EDSS is assessed, or some event or events, unique to NA or EU, that altered the balance of baseline characteristics in the pool of eligible patients.

Primary efficacy endpoint

The primary efficacy endpoint was time to confirmed disease progression, defined as a 1-point increase in EDSS confirmed at two subsequent, consecutive 12-week visits. (A 0.5 point increase was taken as evidence of progression if the baseline EDSS was 6.0 or 6.5.) Because of the requirement for confirmation, the Week 144 visit provided the final opportunity for progression. In the Betaseron groups as well as in the placebo group, 1.9% of subjects had no opportunity for a confirmed progression.

In total, there were 333 progressions in the 939 subjects (an overall progression rate of 35%). There were 106 progressions (34%) recorded in the placebo group, 124 (39%) in the BSA-adjusted Betaseron group, and 103 (32%) in the fixed-dose Betaseron group. Differences among the groups were not statistically significant. The results were generally consistent across sites, and, in particular, between the U.S. and Canadian sites. Thus, with respect to

prevention of disability progression, the study provides no evidence of a Betaseron treatment effect for either regimen. It is noteworthy that in the prior EU study, rates of progression were 55% for the placebo group, versus 47% for the interferon group. For the NA study, therefore, *all* groups performed substantially better than predicted on the basis of the EU study.

Secondary efficacy measures

Amongst the secondary endpoints, the change in EDSS relates to disease progression and is most relevant. On this endpoint, there no significant effect of interferon.

Betaseron treatment was associated with approximately a one-third reduction in the annual relapse rate, an effect that was statistically significant. Of note, there were relatively few exacerbations in these subjects, relative to subjects in the EU study. Specifically, relapse rates in the NA study were ~40% of those reported in the NA study. The apparent magnitude of the interferon treatment effect was consistent between the NA and EU studies, however.

There was clear evidence of a Betaseron treatment effect with respect to the MRI data, although the meaning of this is unknown. Correlations between MRI data and clinical outcomes, as assessed in the EU study, were not strong. In particular, CBER hypothesized that an early response to Betaseron, as evidenced by improvement in MRI parameters, might predict clinically important responsiveness, i.e., prevention of subsequent disability progression. As part of its complete response letter of January 2001, CBER requested exploratory analyses to evaluate this hypothesis, but the sponsor's analyses did not bear out this hypothesis.

Tertiary efficacy measures

The tertiary endpoints on exacerbations were directionally in favor of Betaseron, but of marginal or no statistical significance. The tertiary MRI endpoints were generally supportive of the secondary endpoints. Other measures of effect were less convincing. In particular, there was no clinically meaningful effect of interferon with respect to quality of life neuropsychologic assessment, ESS, or global impression of improvement worsening and stability.

Safety

Interferon was well-tolerated by the majority of subjects. Serious AEs were balanced in frequency between the interferon and placebo arms. The only SAEs identified more frequently in the interferon arms were pneumonia, ataxia, pulmonary embolism, and dehydration.

In general, adverse events were consistent with the known and accepted adverse reactions associated with Betaseron treatment in RRMS. The were exceptions, however, notably rash, bronchitis, thinking abnormal, leg cramps, vaginitis, dry mouth, fungal dermatitis, urinary tract infection, injection site edema, and skin discoloration. In the EU study, rash, abdominal pain, abscess and hypertension, were significantly associated with interferon.

Depression, particularly transient depression soon after institution of treatment, has been a concern with respect to interferons. In the placebo arm, there were no suicide attempts, whereas there were 3 attempted and 1 completed suicide in the Betaseron groups. There were no statistically significant differences between either treatment group and the placebo group in the mean BDI score change from baseline to any time point. The percentages of subjects in the three groups with moderate or severe depression at any time during the study were also similar across all groups, and antidepressant use was similar among the treatment groups. Thus, although indirect assessments were not consistent with Betaseron-induced depression, the

"harder" evidence (i.e., suicide attempts) continues to raise concerns, and suggest that the labeling warning regarding depression and suicide should be maintained (and revised to include the data from the NA and EU studies).

Conclusions

This, the North American multicenter study of Betaseron in SPMS, failed on its primary disability endpoint, and does not corroborate the results of the similar EU study of Betaseron in SPMS (Study ME 93079). Though the sponsor's exploratory analyses suggest that patients with more active disease derive treatment benefit (patients with recent exacerbations or changes in EDSS score, and/or enhancing MRI lesions), the data are not compelling. For the NA dataset, the net treatment effect is zero, and acceptance of the premise that a particular subgroup receives benefit from Betaseron requires acceptance of the converse premise - - that is, that the complementary subgroup is actually *harmed* by Betaseron. The study does not support the conclusion that Betaseron treatment delays progression of disability in any identifiable population or subpopulation of patients. Moreover, no factors were identified that could be used to *define* a subgroup of patients showing decreased progression of disability in response to Betaseron.

In reconciling these findings with the results of the EU study, it is worthwhile to remember that the EU study, though statistically persuasive, demonstrated only a small treatment effect. The sponsor's premise, that patients with significant recent disease activity receive benefit from Betaseron with respect to prevention of disability progression, is an interesting hypothesis that may merit additional study.

Recommendations

Overall, there is only equivocal evidence of Betaseron's effectiveness in preventing accumulation of disability in SPMS, and the sponsor's proposed expansion of the indication statement to include the claim --" should not be allowed.

With respect to MS exacerbations, the applicant has proposed a change in the indication statement to------------------ ------------------ ---
----

-- --- -------------------------
-

Thus, the study provides evidence that Betaseron reduces the frequency but not the severity of attacks. Of note, however, the favorable risk benefit relation was evident in SPMS patients, and the indicated patient population could be expanded to encompass "relapsing forms" of MS.

Overall, an appropriate indication statement would be: "Betaseron is indicated for the treatment of relapsing forms of MS, to reduce the frequency of exacerbations." If the results and/or data from the EU SPMS study are included in labeling, the corresponding information from the NA SPMS should be presented for balance, along with language to provide an appropriate and cogent interpretation of the conflicting study results.